

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 09:04:55 ; Search time 5325 Seconds  
(without alignments)  
10670.933 Million cell updates/sec

Title: US-10-081-739A-1

Perfect score: 1311

Sequence: 1 atggccaagtactccgagct.....gctactcgggggtgggctga 1311

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sv.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sv.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1311	100.0	1311	12	AF504065	AF504065 Synthetic
2	1102.6	84.1	1299	13	AF504064	AF504064 Unculture
3	1085	82.8	1386	1	AF504062	AF504062 Thermococ
4	1040.8	79.4	1386	1	AF504063	AF504063 Thermococ
5	1003.4	76.5	2705	1	AF068255	AF068255 Thermococ
6	991.2	75.6	2179	1	D83793	D83793 Pyrococcus
7	989.6	75.5	2179	6	E13334	E13334 gDNA encodi
8	963.6	73.5	2372	1	AF017454	AF017454 Thermococ
9	792.8	60.5	1480	1	AF177906	AF177906 Thermococ
10	792.8	60.5	1561	1	AF240464	AF240464 Pyrococu
11	792.8	60.5	1740	1	FF096622	U96822 Pyrococcus
12	792.8	60.5	2134	6	E09372	E09372 Heat-resist
13	792.8	60.5	2627	1	AF001268	AF001268 Pyrococu
14	792.8	60.5	3519	6	BD012151	BD012151 Polypepti
15	792.8	60.5	10643	1	AE010170	AE010170 Pyrococu
16	768.2	58.6	1206	6	E14254	E14254 Thermococu
17	146.4	11.2	221	1	AF068256	AF068256 Thermococ
18	87	6.6	12635	1	AE009342	AE009342 Agrobacte
19	87	6.6	13095	1	AE008292	AE008292 Agrobacte
20	85.2	6.5	245050	1	AL627272	AL627272 Salmonell
21	85.2	6.5	300247	1	AE016837	AE016837 Salmonell
22	83.6	6.4	1700	1	STVF7AA	L01643 S.typhimuri
23	83.6	6.4	24156	1	AE008787	AE008787 Salmonell
24	82	6.3	334520	1	AP003588	AP003588 Nostoc sp
25	79.6	6.1	1458	6	AR175875	AR175875 Sequence
26	79.6	6.1	1458	6	BD249239	BD249239 Polypepti
27	79.6	6.1	1458	6	AR401551	AR401551 Sequence
28	79.6	6.1	1458	6	AX037042	AX037042 Sequence
29	78.4	6.0	1539	1	AF438149	AF438149 Bacillus
30	76.8	5.9	1449	6	A23402	A23402 B.lichenifo
31	76.8	5.9	1449	6	A27772	A27772 Alpha amyla
32	76.8	5.9	1452	6	AX705190	AX705190 Sequence
33	76.8	5.9	1536	6	E01158	E01158 DNA sequenc
34	76.8	5.9	1539	6	E12201	E12201 DNA encodin
35	76.8	5.9	1777	6	A17930	A17930 Alpha amyla
36	76.8	5.9	1777	6	I24553	I24553 Sequence 33
37	76.8	5.9	1777	6	I83686	I83686 Sequence 16
38	76.8	5.9	1948	1	BACAMYABL	M38570 B.lichenifo
39	76.8	5.9	1948	1	BLAAMYLG	X03236 Bacillus 11
40	76.8	5.9	1968	1	BACAMYS	M13256 B.lichenifo
41	76.8	5.9	1968	6	A91241	A91241 Sequence 33
42	76.8	5.9	1968	6	AR000567	AR000567 Sequence
43	76.8	5.9	1968	6	AR012365	AR012365 Sequence
44	76.8	5.9	1968	6	AR049549	AR049549 Sequence
45	76.8	5.9	1968	6	AR065754	AR065754 Sequence

# ALIGNMENTS

RESULT 1	AF504065	1311 bp	DNA	linear	SYN 15-JUL-2002
LOCUS	AF504065				
DEFINITION	Synthetic construct alpha-amylase precursor (BD5088) gene, complete cds.				
ACCESSION	AF504065				
VERSION	AF504065.1	GI:21327000			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE	1 (bases 1 to 1311)				
AUTHORS	Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabell,M., Lam,D., Macomber,J., Short,J.M., Robertson,D.E. and Miller,C.				
TITLE	A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY				



Qy	619	TGGGTCGTCAAGACTGGCTGTAACCTGGTGGGAGGCTGGGCGGTTCGGAGAGTACTGGAC	678
Db	607	TGGGTCGTCAAGACTGGCTGTAACCTGGTGGGAGGCTGGGCGGTTCGGAGAGTACTGGAC	666
Qy	679	ACCAAGCTCGACCGCTGTTCTCAACTGGGATACTCGAGCGGTCGCCAGGTCCTTTGACTTC	738

Db	667	ACCAAGCTGAGCGCTGTTCTTCAACTGGGATACTCGAGCGGTGCCAAGGTCCTTTGACTTC	726
Qy	739	GCCTCTACTACAAGATGGATGAGGGCTTTGACAACAAAACATTCCAGCGCTCGTCTCT	798
Db	727	GCCTCTACTACAAGATGAGCGGGCTTCGATACAACAACATTCGCGCCCTGGTGGAC	786

QY	799	GCCCTTTGAGAAACGGCGCAGACTGTGTGTC	TCCCGGGAACCGGTTCAAGGCGGTAAACCTTTGTA	858
Db	787	GCCTCAGATACGGTCAGACAGTGGTCAGCGCGACCCGTTCAAGGCTGTGACGTTTTGTA	846	
QY	859	GCAAACCGACGACACCGATATAATTCTGGAAACAAGTATCCAGCCTACGGGTTCATCTCTCAC	918	
Db	847	GCCAACCGACGATACCGACATAAATCTGGAAACAAGTATCCAGCCTACGGGTTCATCTCTCAC	906	
QY	919	TACGAGGGCCAGCGCGACAATATTCTACCGGCACTACGAGGAGTGGCTCAACAAGGATAAG	978	
Db	907	TACGAGGCCAGCGGACATAATTCTACCGGCACTACGAGGAGTGGCTCAACAAGGATAAG	966	
QY	979	CTCAAGAACCCTCATCTGGATACATATGAGAACCTCGCCGGAGGAGACCGACATAGTCTAC	1038	
Db	967	CTCAAGAACCCTCATCTGGATACATGACAACTCGCCGGAGGAGACGTCATCATCGTTTAC	1026	

	1039	TACGATAACGATGAATCATCTTCTCGTCAGAAACGGCTTACGGGACAAAGCCGGGGCTTATA	1098
Qy			
	1027	TACGACAACGACGAGCTGATATTCTGTGAGAAACGGCTTACGGAAGCAAGCCGGGACTGATA	1086
Db			

Qy	1159	GGCGGTGCATCCACGAGTACTGTGTAACCTCGAGCGCTGGGTAGACAAGTACGTCCTAC	1218
Db	1147	GGCTCGTGTGCATACACGAGTACACCGGCATCTCGGCGGCTGGGTGACACAGTGGGTGGAC	1206
Qy	1219	TCAAGCGGCTGGGTCTATCTCGAAGTCCAGCTTACGACCTCGCCACCGGCAGTATGGC	1278
Db	1207	TCAAGCGGCTGGGTCTACCTCGAGGCTCTTGCCACACCGCGCCACCGGCAGTACGGC	1266
Qy	1279	TACTCCGTGTGGAGCTACTCGGGTGGGCTGA	1311
Db	1267	TACTCCGTCTGGAGCTACTCGGTGTGGTGA	1299

RESULT 3	
AF504062	1386 bp DNA linear BCT 15-JUL-2002
LOCUS	
DEFINITION	Thermococcus sp. GU5L5 alpha-amylase precursor (BD5064) gene, complete cds.
ACCESSION	AF504062
VERSION	AF504062.1 GI:21326994
KEYWORDS	
SOURCE	Thermococcus sp. GU5L5
ORGANISM	Thermococcus sp. GU5L5 Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae; Thermococcus. 1 (bases 1 to 1386) Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabell,M., Lam,D., Macomber,J., Short,J.M., Robertson,D.E. and Miller,C. A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVER AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE J. Biol. Chem. 277 (29), 26501-26507 (2002)
JOURNAL	11994309
PUBMED	2 (bases 1 to 1386) Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabell,M., Lam,D., Macomber,J., Short,J.M., Robertson,D.E. and Miller,C. Direct Submission Submitted (18-APR-2002) Bioinformatics, Diversa Corporation, 4955
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

Directors Place, San Diego, CA 92121, USA

# FEATURES

Location/Qualifiers  
1..1386  
/organism="Thermococcus sp. GU515"  
/mol\_type="genomic DNA"  
/strain="GU515"  
/db\_xref="taxon:195100"  
1..1386  
/gene="BD5064"  
1..1386  
/gene="BD5064"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-amylase precursor"  
/protein\_id="AAM48112.1"  
/db\_xref="GI:21326995"  
/translation="MKPAKLVFLVLSILAGLYAQPAGAAKYLEBEGGVIMQFYW  
DVPSSGIWDDTIRQKIPYDAGISAIWIPPSKGMGGAYSMGYDPYDFDLGEYDQK  
GTVEYRFGSKDELVINHTAHAYGIVKIVADIINHRAGDLEWNPVNDYTWTFPSKV  
ASGKYTANYLDLHPNVEKCDDEGTGGPDIAHEKSWDOYWLWASNEVYAAVLRISGV  
DAWRFDYKYGAVVVKDMLDNWGWAGYWDVNDALLNWAYSSDKAVDFPLAYK  
MDAAFDNKIPALVELKNGTGVUSDDPEKAVTFVANHDTDLIWNKYPAAYAPILTYEG  
QPTIFRDRYBEWLNDRKMLIWIHDLAGSDIDIVYDNDLIIIVRNGYDGKPLGIT  
YINLSSKAGRWYVPKFAAGACIHEYTNLGNLWVNDKWDSSGWVYLEAPADHPANGYY  
GYSVWSYCGVG"  
1..78  
/gene="BD5064"

sig\_peptide

## ORIGIN

Query Match 82.8%; Score 1085; DB 1; Length 1386;  
Best Local Similarity 89.3%; Pred. No. 1.2e-234;  
Matches 1169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3 GGCCTAGTACTCGAGCTGGAAGAAGCGCGGTCTAATGACAGGCGTCTACTGGAGCT 62  
78 GGCCTAGTACTCGAGCTGGAAGAAGCGCGGTCTAATGACAGGCGTCTACTGGAGCT 137  
63 GCCTTCAGGAGGAATATGTGGACACAATACGGCAGAGATACCGGAGTGGTACGATGC 122  
138 GCCTTCAGGAGGAATATGTGGACACAATACGGCAGAGATACCGGAGTGGTACGATGC 197  
123 CGGAATCTCGCAATATGATATCCCGCGGAGCAGGCGATGGCGGCGCTATTCCGAT 182  
198 CGGAATCTCGCAATATGATATCCCGCGGAGCAGGCGATGGCGGCGCTATTCCGAT 257  
183 GGGCTACGACCCCTACGACTCTTTTGAACCTCGGTGAGTACGACCAAGAGGAAACGGTAGA 242  
258 GGGCTACGACCCCTACGACTCTTTTGAACCTCGGTGAGTACGACCAAGAGGAAACGGTAGA 317  
243 GAGCGCTTTGGCTCAAGCAGGAGCTCGTGAACATGATATAACACCGCGCCGCTATGG 302  
318 GAGCGCTTTGGCTCAAGCAGGAGCTCGTGAACATGATATAACACCGCGCCGCTATGG 377  
303 CATGAAGGTAATAGCGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGGAA 362  
378 CATGAAGGTCATCGCAGACATAGTATCAACACCGCGCGCGGTGACCTGGAGTGGAA 437  
363 CCCCTTCGTGACGACTATACCTGACCGGACTTCTCAAGGTTCGGTGGGTAATAACAC 422  
438 CCCCTTCGTGACGACTATACCTGACCGGACTTCTCAAGGTTCGGTGGGTAATAACAC 497  
423 GGCCTACTACCTCGACTTCCACCCGAAACGAGCTCCATGCGGCGGATTCGGAAACATTTGG 482  
498 GGCCTACTACCTCGACTTCCACCCGAAACGAGCTCCATGCGGCGGATTCGGAAACATTTGG 557  
483 AGGCTATCCCGACATATGCGACGACAGAGTGGGACGAGTACTGGCTCTGGGCGAGCA 542  
558 AGGCTATCCCGACATATGCGACGACAGAGTGGGACGAGTACTGGCTCTGGGCGAGCA 617  
543 GAGAGAGTACGCGGATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAGCT 602  
618 CGAGAGTACGCGGATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAGCT 677

QY 603 CAAGGGCTATGCTCCCTGGGTGCTCAAGGAGTGGTGAACCTGCTGGGAGGCTGGGCGGT 662  
Db 678 CAAGGGCTATGCTCCCTGGGTGCTCAAGGAGTGGTGAACCTGCTGGGAGGCTGGGCGGT 737  
QY 663 TGGAGAGTACTTGGGACACCAACGTCGACCGTGTCTCAACTGGGACATATCTCGAGCGGTGC 722  
Db 738 CGGGAGTACTTGGGACACCAACGTTGATGCTCACTCACTGGGCTACTCGAGCGATGC 797  
QY 723 CAAGGTCTTTGACTTCCGCTCTACTACAAGATGAGTGGGCTTTGACAAACAAACAT 782  
Db 798 AAAAGTCTTCCGCTTCCGCTCTACTACAAGATGAGTGGGCTTTGACAAACAAACAT 857  
QY 783 TCCAGCGCTGCTCTCTCCCTTCAGAACGCCAGACATGTTGTCTCCGCGGACCGCTTCA 842  
Db 858 TCCGCGACTCTGTCGAGGCCCTCAAGAACGGGGGACAGTCTGAGCGCGGACCGCTTTAA 917  
QY 843 GGCCTTAACCTTTGTAGCAAAACCAACGACCGATATAATCTGGAAACAAGTATCCAGCTA 902  
Db 918 GGCCTTAACCTTTGTAGCAAAACCAACGACCGATATAATCTGGAAACAAGTATCCAGCTA 977  
QY 903 CGGTTTCATCTCCTACCTACGAGGGCCGACCAATATTTCTACCGGACTACGAGGAGTG 962  
Db 978 GGCCTTCATCTCCTACCTACGAGGGCCGACCAATATTTCTACCGGACTACGAGGAGTG 1037  
QY 963 GCTCAACAGGATAAGCTCAAGAACCTCATCTGATACATAGAACTCCCGCGGAGGAG 1022  
Db 1038 GCTCAACAGGATAAGCTCAAGAACCTCATCTGATACATAGAACTCCCGCGGAGGAG 1097  
QY 1023 CACGACATAGTCTACTACGATTAACGATGAACATCTCTCTGAGGACCGGCTACGGGA 1082  
Db 1098 CACGACATAGTCTACTACGATTAACGATGAACATCTCTCTGAGGACCGGCTACGGGA 1157  
QY 1083 CAAGCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGAGGTTGGTTA 1142  
Db 1158 CAAGCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGAGGTTGGTTA 1217  
QY 1143 TGTGCGAAGTTCGCGGGCGGTGCTATCAAGAGTATCTGTAACCTCGAGGCTGGGT 1202  
Db 1218 CGTTCGCAAGTTCGCGGGCGGTGCTATCAAGAGTATCTGTAACCTCGAGGCTGGGT 1277  
QY 1203 AGCAAGTACTGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGC 1262  
Db 1278 GGAAGTGGTGGATCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGC 1337  
QY 1263 CAAGCGGAGTATGGTCTACTCCGCTTGGAGCTACTCGGGTGGGCTGA 1311  
Db 1338 CAAGCGGATTTACGGCTACTCCGCTTGGAGCTACTCGGGTGGGCTGA 1386

## RESULT 4

AF504063  
LOCUS  
DEFINITION Thermococcus sp. 'AEP11 la' alpha-amylase precursor (BD5031) gene, complete cds.  
ACCESSION AF504063  
VERSION AF504063.1 GI:21326996  
KEYWORDS  
SOURCE ORGANISM Thermococcus sp. 'AEP11 la'  
Thermococcus sp. 'AEP11 la'  
Archaea; Euryarchaeota; Thermococci; Thermococcales;  
Thermococcaceae; Thermococcus.  
1 (bases 1 to 1386)  
Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabelli,M., Lam,D.,  
Macomber,J., Short,J.M., Robertson,D.B. and Miller,C.  
A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE  
J. Biol. Chem. 277 (29), 26501-26507 (2002)  
11994309  
2 (bases 1 to 1386)  
Richardson,T.H., Tan,X., Frey,G., Callen,W., Cabelli,M., Lam,D.,  
Macomber,J., Short,J.M., Robertson,D.B. and Miller,C.  
Direct Submission  
Submitted (18-APR-2002) Bioinformatics, Diversa Corporation, 4955



Directors Place, San Diego, CA 92121, USA

# FEATURES

## source

1..1386  
/organism="Thermococcus sp. 'ABPII 1a'"  
/mol\_type="genomic DNA"  
/strain="ABPII 1a"  
/db\_xref="taxon:195101"

## gene

1..1386  
/gene="BD5031"

## CDS

1..1386  
/gene="BD5031"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-amylase precursor"  
/protein\_id="AAM48113.1"  
/db\_xref="GI:2132697"  
/translation="MKKEVALFIMFVSVMAVVAQSAAYSELEGGVIMQAFYV  
DYPGGIWDITRISXIPWEGYSAIWIPIPAKSMGSGYDIPYDFTDGLGYNOK  
GTIEFRGSKBLINIMTAAYGKIKIADIVINHRAGGDLWNPFDVDTWTFDSKV  
ASGKYANLDFPHNEKCCDEGTGGPDIHAKESMDQHWLWASDRSYAAYLBSIGV  
DAMREDYKGVGMVVDKMLNMGWMAVEYWDITNVDALLNWAYSSGAKUPDFPLYK  
XDEAFDNKNIIPALVSALQNGQTVRSRDFKAVTFVANHDTDIIWNKYLAFILITYSG  
QVPIPYRDEBWLKNDRLNLIWIHDLHAGGSTIVYDSDMIFVRNGYSGKPLGIT  
YINLGSSKVGWVYVPKFPAGACIHEYTGNLGQWVDKYYVSSGWVYLEAPAYDPANGQY  
GYSWMSYCGVG"  
1..78  
/sig\_peptide  
/gene="BD5031"

## ORIGIN

Query Match 79.4%; Score 1040.8; DB 1; Length 1386;

Best Local Similarity 87.2%; Pred. No. 1.2e-224;

Matches 1141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

QY 4 GCCAAGTACTCGAGCTGGAAGAGCGGGCTCAATGCGAGGCGTCTTACTGGGAGTG 63
Db 79 GCAAGTATTCGAGCTCGAAGAGCGGGCTTAAATGCGAGGCGCTTCTACTGGGAGTG 138
QY 64 CTTTCAGGAGGAATATGTTGGGACACAATACGACAGAAAGATACCGAGTGTGTAGCATGCC 123
Db 139 CCAGTGTGAGGAATCTGTGTGGGACACCATCAGGACCAAGATACCGAGTGTGTAGCGGG 198
QY 124 GGAATCTCGGCAATATGATTCCTCCCGCGGAGCAGAGGCGGCGGCGCTATTTCGATG 183
Db 199 GGAATATCGGCAATATGATTCCTCCCGCGGAGCAGAGGCGGCGGCGGCTATTTCGATG 258
QY 184 GGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAAGCGTAGAG 243
Db 259 GGTACGATCCCTACGATTTCTTTGACCTCGGCGAGTACACCAAGAGGAAAGCGTAGAG 318
QY 244 ACGCGCTTTGGCTCCAGCAGAGCTCGTGAAACATGATATAACACCGCCCGCTATGGC 303
Db 319 ACGCGCTTTGGCTCCAGCAGAGCTCATCAATATGATATAACACCGCCCGCTATGGC 378
QY 304 ATGAAGGTAATAGCGATATAGTATCAACACCGCGCGCGGCTGACCTGGAGTGGAC 363
Db 379 ATAAAGGTCATAGCGGATCGTCTAAACACCGCGCGCGGAGAGCTCGAGTGGAC 438
QY 364 CCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAAGGTGCGGTTCGGGTAAATACAG 423
Db 439 CGCTTCGTGGGAGTACACCTGGAGCGACTTCTCAAAAGGTGCGGTTCGGGTAAATACAG 498
QY 424 GCCAACTACCTCGACTTCACCCGAGCAGCTCCATGCGGCGGATTCGGAACATTTGGA 483
Db 499 GCCAACTACCTCGACTTCACCCGAGCAGCTCCATGCGGCGGATTCGGAACATTTGGA 558
QY 484 GGCTATCCGACATATGACACAAAGAGCTGGGACCACTACTGGCTCTGGGCGGACGAG 543
Db 559 GGCTTCCGACATATGACACAAAGAGCTGGGACCACTACTGGCTCTGGGCGGACGAG 618
QY 544 GAGAGCTACGGCGGATATCTCAGGAGCATTCGCGATTCGATCGCTCGGCTTCGATCGTC 603
Db 619 GAGAGCTACGGCGGCTACCTAAGGAGCATCGGCGTTGATGCTCGGCGCTTGACTAGTG 678

```

```

QY 604 AAGGGCTATCTCCCTGGGTCTGTCAGGAGCTGCTGAACTGTGGGAGAGCTCGGCGGTT 663
Db 679 AAGGGCTATCGGAGCTGGGTCTGTCAGGAGCTGCTGAACTGTGGGAGAGCTCGGCGGTT 738
QY 664 GGAGAGTACTGGGACACCAACGTCGAGCGTGTCTCAACTGGGACATCTCGAGCGGTGCC 723
Db 739 GCGAGTACTGGGACACCAACGTTGATGACTCTCACTGGGCTTACTCGAGCGGCGCC 798
QY 724 AAGGTCTTTGACTTCGCGCTTACTACTCAAGATGATGAGCGCTTTGACAAACAAACATT 783
Db 799 AAGGTCTTCGACTTCGCGCTTACTACTCAAGATGATGAGCGCTTTGACAAACAAACATT 858
QY 784 CCAGCGCTGCTCTGCGCTTCTGAGAGCGGCGGAGCTGTTGCTCCGCGGACCGTTCAG 843
Db 859 CCAGCGCTGCTCTGCGCTTCTGAGAGCGGCGGAGCTGTTGCTCCGCGGACCGTTCAG 918
QY 844 GCGTAACTCTTTAGCAAAACCAACGACACCGATATATCTGGAACAAAGTATCCAGGCTAC 903
Db 919 GCGTAACTCTTTAGCAAAACCAACGACACCGATATATCTGGAACAAAGTATCCAGTAT 978
QY 904 GCGTCACTCTCACTACGAGGCGGCGGACCAATATTTCTACCGGACTACGAGGAGTGG 963
Db 979 GCTTTTATCTCTCACTACGAGGCGGCGGCGGCTCATATTTCTACCGGACTACGAGGAGTGG 1038
QY 964 CTCAACAGGATTAAGCTCAAGAACCTCATCTGATGATACATGAGAACCTCGCGGAGGAAGC 1023
Db 1039 CTCAACAGGATTAAGCTCAAGAACCTCATCTGATGATACATGAGAACCTCGCGGAGGAAGC 1098
QY 1024 ACCGACATAGTCTACTACGATTAACGATGAACTCATCTTCTGAGAACCGGCTACGGGAC 1083
Db 1099 ACGAGCATAGTCTACTACGAGGCGGCGGCTCATCTTCTGAGAACCGGCTATGGAAGC 1158
QY 1084 AAGCGGGGCTTAACTACATCAACCTAGGCTGAGCAAGCGGCGGAGGCTGGGTTAT 1143
Db 1159 AAGCTTGGGCTTAACTTACATCAACCTGCGGCTGAGCAAGGTTGGAAGGTTGGGTTAT 1218
QY 1144 GTCCGGAAGTTCGCGGCGGCTGATCCAGAGTACTTGGTAACTCGGAGGCTGGGTA 1203
Db 1219 GTCCGGAAGTTCGCGGCGGCTGATCCAGAGTACTTGGTAACTCGGAGGCTGGGTA 1278
QY 1204 GACAGTACTGCTACTCAAGCGGCTGGGCTTACTTCAAGAGTCCAGCTTACGACCTGCC 1263
Db 1279 GACAGTACTGCTACTCAAGCGGCTGGGCTTACTTCAAGAGTCCAGCTTACGACCTGCC 1338
QY 1264 AAGCGGAGTATGGCTACTCGGTGAGCTACTCGGCGGCTGGGCTGA 1311
Db 1339 AAGCGGAGTATGGCTACTCGGTGAGCTACTCGGTGAGCTATGGGCTGA 1386

```

RESULT 5  
AF068255  
LOCUS  
DEFINITION  
Thermococcus hydrothermalis strain AL662 alpha-amylase (amy) gene,  
complete cds.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AF068255.1 GI:4049917  
Thermococcus hydrothermalis  
Thermococcus hydrothermalis  
Archaea: Euryarchaeota; Thermococci; Thermococcales;  
Thermococcaceae; Thermococcus.

REFERENCE  
1 (bases 1 to 2705)

AUTHORS  
Leveque,E., Belarbi,A. and Haye,B.

JOURNAL  
Patent: France 98.05655 05-MAY-1998;

REFERENCE  
2 (bases 1 to 2705)

AUTHORS  
Leveque,E., Haye,B. and Belarbi,A.

TITLE  
Cloning and expression of an alpha-amylase encoding gene from the  
hyperthermophilic archaeobacterium Thermococcus hydrothermalis and  
biochemical characterisation of the recombinant enzyme

JOURNAL  
FEMS Microbiol. Lett. 186 (1), 67-71 (2000)

MEDLINE  
20243462

PUBMED  
10779714

REFERENCE  
3 (bases 1 to 2705)

AUTHORS Leveque, E., Hays, B. and Belarbi, A.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-MAY-1998) Laboratory of General and Molecular Microbiology, UFR Sciences, University of Reims-Champagne-Ardenne, Moulin de la Housse, BP 1039, Reims Cedex 2 51687, France

## FEATURES

Location/Qualifiers  
 1..2705  
 /organism="thermococcus hydrothermalis"  
 /mol\_type="genomic DNA"  
 /strain="AL562"  
 /db\_xref="taxon:46539"  
 424..1797  
 /gene="amy"  
 424..1797  
 /gene="amy"  
 /EC\_number="3.2.1.1"  
 /codon\_start=1  
 /transl\_table=11  
 /product="alpha-amylase"  
 /protein\_id="AAC97877.1"  
 /db\_xref="GI:4049918"  
 /translation="MARKVIVALLVPLVLSVAVPAKATLENGGVIMQAFYWDVPG  
 GIVWDITIAQKPDWASAGISAIWIPVPSKMGSGYSDYDFDILGEYKQSGVE  
 TFGSKBELVNTIAHAKMVIADIVINHRAGDLEWNPFTNSYTWDPSKVASGK  
 YTYLDFHPLHAGDSGTFGYDPICHDKSWDQHLWASNESYAAVLRISIGIDAWR  
 EDVYKGA PWVVKWLNRRGWAGVEYDNTVDALLSWAYDSKAVFDFLYKXDEA  
 FNNNTPALVDKLNKGSTVSDPDKATVFVANHDTNIIWNKYPAVAFILYVEGPAI  
 FRDYEWLNKRLNLIWIHDLHAGGSDIILYDSDELIIVRNGYGGKPLIYINL  
 GSKAGRWVYVPKFGASCIHEYTNLGGWIDRWVSSGRVYLEAPHPANGQYGSV  
 WSYCGVG"

## ORIGIN

Query Match 76.5%; Score 1003.4; DB 1; Length 2705;  
 Best Local Similarity 86.0%; Pred. No. 3 4e-216;  
 Matches 1112; Conservative 0; Mismatches 181; Indels 0; Gaps 0;  
 19 CTGAAAGAGCGGGGTCTAATGCAAGGCTTCTACTGGAGCGTGCCTTCAGAGGAATA 78  
 505 CTTGAGAACCGCGCGCTCATATGCAAGGCTTCTACTGGAGCGTGCCTTCAGAGGAATC 564  
 79 TGGTGGGACACATACCGGACAGATACCGGAGTGTACGATGCGGAATCTCCGCAATA 138  
 565 TGGTGGGACACATACCGGACAGATACCGGAGTGTACGATGCGGAATCTCCGCAATA 624  
 139 TGGATTCCTCCCGGAGTGGGCGGCGCTTATTCGATGGGCTACGACCCCTAC 198  
 625 TGGATTCCTCCCGGAGTGGGCGGCGCTTATTCGATGGGCTACGACCCCTAC 684  
 199 GACTTCCTTCGAGTGGTACGACGACGAGGAGACCGTAGAGCGGCTTGGCTCC 258  
 685 GATTTCCTTCGAGTGGTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 744  
 259 AAGCAGGAGCTCGTGAACATGATTAACACCGCCCGCTATGGCATGAAGTAAATAGCC 318  
 745 AAGCAGGAGCTCGTGAACATGATTAACACCGCCCGCTATGGCATGAAGTAAATAGCC 804  
 319 GATATAGTCATCAACACCGCGGCGGTGACCTGGAGTGGAAACCCCTTCGTGAAACGAC 378  
 805 GACATAGTCATCAACACCGCGGCGGTGACCTGGAGTGGAAACCCCTTCGTGAAACGAC 864  
 379 TATACCTGGACCGGCTTCTCAAGGTCGCTCGGCTGAATATACACGCGCAACTACCTCGAC 438  
 865 TACACCTGGACCGGCTTCTCAAGGTCGCTCGGCTGAATATACACGCGCAACTACCTCGAC 924  
 439 TTCCACCGGACGAGCTTCCATGGGCGGATTCGGAACATTTGGAGGCTATCCGACATA 498  
 925 TTCCACCGGACGAGCTTCCATGGGCGGATTCGGAACATTTGGAGGCTATCCGACATA 984  
 499 TGCACACGAGAGCTGGGACGAGTACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 558  
 985 TGCACACGAGAGCTGGGACGAGTACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1044  
 559 TATCTCAGGAGCATCGGCATCGATCGCTCGGCTTCGACTACGTCAAGGCGCTATGCTCCC 618

Db 1045 TACCTCCGGAGCATGGCATCGAGCGCTGGCGCTTCGACTACGTCAAGGGCTACGCTCCC 1104  
 Qy 619 TGGGTCTCTAAGGACGTGGCTGAACCTGGTGGGAGGCTGGGCGGTTGGAGAGTACTGGGAC 678  
 Db 1105 TGGGTCTCTAAGGACGTGGCTGAACCTGGTGGGAGGCTGGGCGGTTGGAGAGTACTGGGAC 1164  
 Qy 679 ACCAAGCTCGACGCTGTCTCAACTGGGCATCTCGAGGCGGTGCCAAGGTCTTTGACTTC 738  
 Db 1165 ACCAAGCTCGATGCACTCTGAGCTGGGCTACGACGCGGTCTAAGTCTTCGACTTC 1224  
 Qy 739 GCCCTCTACTAAGATGGATGAGGCTTTGACAAACAAACAAATTCAGCGCTTCGTCTCT 798  
 Db 1225 CCGCTCTACTAAGATGGATGAGGCTTCGATAACAAACATCCCGCCCTCGTGGAC 1284  
 Qy 799 GCCCTTCAGAACCGGACGCTGTCTCCCGGAGCGGCTTCAGGCGGTTCAGGCGCTTGTGA 858  
 Db 1285 GCCCTCAAGAACCGGAGCGGCTGTGACGCGGAGCCCGTTCAAGCGGTGACCTTCGTT 1344  
 Qy 859 GMAAACCAACGACCGCATATAATCTTGAAACAAGTATCCAGCCTACGCGTTTCATCTCACC 918  
 Db 1345 GCCAACCAGTACCAACATAATCTGGAAACAAGTATCCGCGCTACGCTTCATCTCACC 1404  
 Qy 919 TACGAGGCGGACGACCAATATTTCTACCGGACATACGAGGAGTGGCTCAACAGGATAAG 978  
 Db 1405 TATGAGGACGACGCGGCAATATTTCTACCGGACATACGAGGAGTGGCTCAACAGGACAG 1464  
 Qy 979 CTCAAGAACCTCTCTCGATACATGAGAACCTCGCGGAGGAGACCGACATAGTCTAC 1038  
 Db 1465 CTCAAGAACCTCTCTCGATACATGAGAACCTCGCGGAGGAGACAGACATCATCTAC 1524  
 Qy 1039 TACGATAACGATGAATCTCTCTCGTCAAGAACCGGTACCGGAGACAGCGCGGCTTATA 1098  
 Db 1525 TACGACGACGACGAGCTTATCTCTGAGAAACCGGTACCGGAGACAGCGCGGACTGATA 1584  
 Qy 1099 ACCTACATCAACCTAGGCTCGAGCAGCGGAGGAGTGGGCTTTATGTCCGAAAGTTCGCG 1158  
 Db 1585 ACCTACATCAACCTAGGCTCGAGCAGCGGAGGAGTGGGCTTTATGTCCGAAAGTTCGCG 1644  
 Qy 1159 GGCGCGTGCATCCACGAGTATCTGTAACCTCGGAGGCTGGGTAGACAGTACGCTAC 1218  
 Db 1645 GGCTCGTGCATACACGAGTACACCGGCAACCTCGCGGCTGGATTCACAAAGTGGGTGAC 1704  
 Qy 1219 TCAAGCGGCTGGGTCTATCTCGAAGCTTCAGAGCTTCAGACCTCCGACCGGAGTATGGC 1278  
 Db 1705 TCAAGCGGCTGGGTCTATCTCGAAGCTTCAGAGCTTCAGACCTCCGACCGGAGTATGGC 1764  
 Qy 1279 TACTCCGTGTGGAGCTACTCGCGGCTGGGCTGA 1311  
 Db 1765 TACTCCGTGTGGAGCTACTCGCGGCTGGGCTGA 1797

## RESULT 6

D83793

LOCUS

DREINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

Direct Submission

Submitted (05-MAR-1996)

Yoshihisa Tachibana, Osaka University,

Takagi, M. and Imanaka, T.

Takagi, M. and Imanaka, T.

Takagi, M. and Imanaka, T.

Takagi, M. and Imanaka, T.

Takagi, M. and Imanaka, T.

Takagi, M. and Imanaka, T.

Takagi, M. and Imanaka, T.

Faculty of Engineering, Biotechnology; 2-1, Yamadaoka, Suita, Osaka  
565, Japan (Tel:06-879-7442, Fax:06-879-7448)

## FEATURES

Location/Qualifiers  
1..2179  
/organism="Pyrococcus sp."  
/mol\_type="genomic DNA"  
/strain="KOD1"  
/db\_xref="taxon:33866"  
452..1837  
/gene="Apka"  
452..1837  
/gene="Apka"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-amylase"  
/protein\_id="BAA21130.1"  
/db\_xref="GI:2251108"  
/translation="MKFVLLITMPVSVMAVAOPASAAKYLEEGGVMQAFYV  
DVPAGGIWDTIRSKIPWEYEGISAIHIPPASKMGXAYSMGYDIPDFDLGEYNQK  
GVETRGSGQELINHTAHYGIKVIADIVINHRAGGLEWNPFGDYITWDPKVG  
ASGKYTYLDHPNEVKCDEGTFGGFDIAHEKWDQHLWASDEKAYLRSIGV  
DAWRFYVGYGAWVVDNLNMGWAGVYNDVNDLNNWAYSGKAVDFPFLYK  
MDFAFDNTNIPALVDLQNGTIVSRDPFKAFTFVANHDTDIINWYKPAYAFLITVEG  
QPIFYRDEBWLKDKLNLIIWHDLHLAGSTSIIVYDSDELIFVRNGDSKRPGLIT  
YINLSSKVGWRVYVPKEAGACIHEYTNLGNVVDKVVESSGWVYLEAPYDPAQQY  
GITVNSYCGVG"  
452..529  
/gene="Apka"  
530..1834  
/gene="Apka"  
/product="alpha-amylase"

## sig\_peptide

## mat\_peptide

## ORIGIN

Query Match 75.6%; Score 991.2; DB 1; Length 2179;  
Best Local Similarity 84.9%; Pred. No. 2e-213;  
Matches 1110; Conservative 0; Mismatches 198; Indels 0; Gaps 0;  
4 GCCAAGTACTCCGAGTCGAAAGCGGGGCTCATATGCGAGCGTCTACTCGGAGTG 63  
530 GCAAGTATCCGAACCTCGAAGAGGGCGGCTTATAATGCAAGGCTTCTACTGGAGTT 589  
64 CTTTCAGGAGGAATATGTTGGGACACAAATAGCGGAGAGATACCGGAGTGGTACGATGCC 123  
590 CCACGGGAGGAATCTGGTGGAATACATCAGAGAGCAAGATACCGGAGTGGTACGAGGCT 649  
124 GGAATCTCCGAATATGATATCCCGCGGAGGAGGCAATGCGGCGGCTTATCGATG 183  
650 GGAATCTCCGCAATCTGATTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 709  
184 GGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTAGAG 243  
710 GGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTAGAG 769  
244 ACGGCTTTGGCTCAGAGCAGGAGCTCGTGAAATGATTAACACCGCCCGCCTATGGC 303  
770 ACTCGCTTTGGCTCAAGAGCAGGAGCTTATCAACATGATTAACACCGCCCGCCTATGGC 829  
304 ATGAGGTAATAGCGATATAGTATCAACACCGCGCGCGGCTGACCTGGAGTGAAC 363  
830 ATAAAGGTATAGTGTATATCGTATTAACCGCGCGCGGAGGAGGAGGAGGAGGAGGAGG 889  
364 CCCTTCGTGAACGACTATACCTGGAACCGACTTCTCAAGGTTCGGTTCGGGTAAATACAG 423  
890 CGTTTCGTGGGACTACCTGGAACCGACTTCTCAAGGTTCGGTTCGGGTAAATACAG 949  
424 GCCAATCTCGATTCGATTCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 483  
950 GCCAATCTCGATTCGATTCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1009  
484 GGCTATCCGACATATGCGACGAGAGCTGGGACGAGTACTGGCTCTGGGCGGAGCGAG 543  
1010 GGTTCCTCCAGACATAGCCACGAGAGGAGTGGGACGAGTACTGGCTCTGGGCGGAGCGAG 1069

544 GAGAGCTACGGGCGCATATCTCAGGAGCATCGGATCGATGCTGCTGCGGCTTCGACTACGTC 603  
1070 GAGAGCTACGGGCGCATCTCAGGAGCATCGGCGTGTGATGCTGCGGCTTCGACTACGTA 1129  
604 AAGGCTATGCTCCTCGGTCGTCAAGGACTGGCTGAACTGCTGGGAGGCTGGGCGGTT 663  
1130 AAGGCTACGGAGCGTGGTCTCAAGGACTGGCTCAACTGCTGGGAGGCTGGGCGGTC 1189  
664 GGAGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCATCTCAGAGCGGTGCC 723  
1190 GGTGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCATCTCAGAGCGGCGCC 1249  
724 AAGGCTTTCGATTCGCGCTCTACTCAAGATGGATGAGGCTTTGACAAACAAACATTT 783  
1250 AAGGCTTTCGATTCGCGCTCTACTCAAAATGGACGAGGCTTTGACAAACAAACATTT 1309  
784 CCAGCGCTGCTCTGCGCTTCAGAAACGCGCAGACTGTCTCTCCGCGGACCGGTTCAAG 843  
1310 CGGCTCTGGTTGATGCGCTTCAGAAACGCGGGAACCGTCTCTCGGAGCGGTTCAAG 1369  
844 GCGCTAACCTTTGTAGCAAAACCAACACCGGATATAATCTGGAACAAATATCTCAGCTAC 903  
1370 GCGCTAACCTTTGTAGCAAAACCAACACCGGATATAATCTGGAACAAATATCTCAGCTAC 1429  
904 GCGCTAACCTTCACCTACGAGGCGGCGGACCAATATCTCAGGAGTACGAGGAGTGG 963  
1430 GCTTTCATCTCCTCCTACGAGGCGGCGGCGGCTCATATCTCAGGAGTACGAGGAGTGG 1489  
964 CTCAACAGGATTAAGCTCAAGAACCTCATCTGGATATACGAACTCTCTGTCAGGAGGAGTAC 1023  
1490 CTCAACAGGATTAAGCTTAACAACTAATCTGGAATACGACCACTCGCGGTTGGAAGC 1549  
1024 ACGGATAGTCTACTACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG 1083  
1550 ACGGATAGTCTACTACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG 1609  
1084 AAGCGGCGGCTTATACCTTACATCACTAGGCTCGAGCAAGCGCGGAGTGGGTTTAT 1143  
1610 AAGCGGCGGCTTATACCTTACATCACTAGGCTCGAGCAAGCGCGGAGTGGGTTTAT 1669  
1144 GTGCGGAATTCGCGGCGGCTGATCCAGAGTATCTGGAATCTCGAGGCTGGGTA 1203  
1670 GTGCGGAATTCGCGGCGGCTGATCCAGAGTATCTGGAATCTCGAGGCTGGGTA 1729  
1204 GACAGTACGCTTACTCAAGCGGCTGGGCTCATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
1730 GACAGTACGCTTACTCAAGCGGCTGGGCTCATCTCGAAGCTCCAGCTTACGACCTGCC 1789  
1264 AAGCGGCGGCTTACTCGGCTTACTCGGCTTACTCGGCTTACTCGGCTTACTCGGCTTACT 1311  
1790 AAGCGGCGGCTTACTCGGCTTACTCGGCTTACTCGGCTTACTCGGCTTACTCGGCTTACT 1837

## RESULT 7

E13334  
LOCUS  
DEFINITION  
GDNA encoding alpha-amylase.  
E13334  
ACCESSION  
VERSION  
E13334.1 GI:3252139  
KEYWORDS  
JP1997173077-A/1.  
SOURCE  
unidentified  
ORGANISM  
unclassified.  
REFERENCE  
1 (bases 1 to 2179)  
AUTHORS  
Imanaka, T., Tachibana, Y., Suzuki, Y., Kojima, I. and Uzura, K.  
TITLE  
ULTRATHERMOSTABILE ACIDIC ALPHA-AMYLASE AND DNA FRAGMENT  
CONTAINING GENE FOR PRODUCING THE ALPHA-AMYLASE  
JOURNAL  
Patent: JP 1997173077-A 1 08-JUL-1997;  
IMANAKA TADAYUKI, NAGASE SEIKAGAKU KOGYO KK  
COMMENT  
CS Unclassified  
PN JP 1997173077-A/1  
PD 08-JUL-1997  
PF 19-JUL-1996 JP 1996191138

PR 20-JUL-1995 JP 95P 216454  
 PI IMANAKA TADAYUKI, TACHIBANA YOSHINAGA, SUZUKI YUJI, PI  
 KOJIMA IWAO,  
 PC C12N15/09, C07H21/04, C12N1/21, C12N9/30, C12P19/14, (C12N15/09, PC  
 C12R1:01),  
 CC (C12N1/21, C12R1:19), (C12N9/30, C12R1:19);  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No; Location/Qualifiers

FT source 1. .2179  
 FT CDS /organism='Unclassified'  
 FT /strain='KOD-1',  
 FT /product='alpha-amylase'.  
 FT Location/Qualifiers

1. .2179  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

Query Match 75.5%; Score 989.6; DB 6; Length 2179;  
 Best Local Similarity 84.8%; Pred. No. 4.5e-213;  
 Matches 1109; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

4 GCCAGTACTCCGAGCTGGAAGAGGGCGGTCATTAATGCAGGCGTTCTACTGGACGTG 63  
 530 GCAAGATATTCGCACTCCGAGAGCGCGGCTTAAATGCAGGCGTTCTACTGGACGT 589

64 CTTTCAGGAGGATATGTTGGGACACATACCGGAGAGATACCGAGTGGTACGATGCC 123  
 590 CCAGCGGAGGATCTGTTGGATACATCAGAACAGATACCGAGTGGTACGAGGCT 649

124 GGAATCCGCAATATGATATCCCGCGGAGCAAGGGGATGGCGGCGCTTATTCGATG 183  
 650 GGAATCCGCGCATCTGATATCCCGCGGAGCAAGGGGATGGCGGCGCTTATTCGATG 709

184 GGTACGACCCCTACGACTCTTTGACCTCGTGTAGTACGACCAAGAGGAGCGTAGAG 243  
 710 GGTACGACCCCTACGACTCTTTGACCTCGTGTAGTACGACCAAGAGGAGCGTAGAG 769

244 ACGCGCTTTGGTCCAGCAGAGCTCGTGAAATGATAAACAACCGCCACGCTTATGGC 303  
 770 ACTCGCTTTGGTCCAGCAGAGCTTATCAACATGATAAACAACCGCCACGCTTATGGC 829

304 ATGAGGTATACCGATATAGTATCAACACCGCGCGGCGGCTGACCTGGAGTGGAC 363  
 830 ATGAGGTATAGTATAGTATCAACACCGCGCGGCGGCTGACCTGGAGTGGAC 889

364 CCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGAGTGGCTCGGCTGAATACACG 423  
 890 CCCTTCGTGGGACTATACCTGGACCGACTTCTCAAGAGTGGCTCGGCTGAATACACG 949

424 GCACTACTCGACTTCAACCGCAACGATCCATGCGGGGATTCGGAACATTTGGA 483  
 950 GCACTACTCGACTTCAACCGCAACGATCCATGCGGGGATTCGGAACATTTGGA 1009

484 GGTATCCGACATATGCAACAGAGCTGGGACGACTGGCTCTGGCGGAGCGGAC 543  
 1010 GGTATCCGACATATGCAACAGAGCTGGGACGACTGGCTCTGGCGGAGCGGAC 1069

544 GAGAGTACCGGCAATATTCAGAGGATTCGAGCATGATGCGGCTTCGATACGTC 603  
 1070 GAGAGTACCGGCAATATTCAGAGGATTCGAGCATGATGCGGCTTCGATACGTC 1129

604 AAGGGCTATGCTCCCTGGGCTCGTCAAGGACTGGCTGAATGTTGGGAGGCTGGCGGTT 663  
 1130 AAGGGCTACGGAGCGTGGGCTCGTCAAGGACTGGCTGAATGTTGGGAGGCTGGCGGTT 1189

## ORIGIN

Query Match 75.5%; Score 989.6; DB 6; Length 2179;  
 Best Local Similarity 84.8%; Pred. No. 4.5e-213;  
 Matches 1109; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 4 GCCAGTACTCCGAGCTGGAAGAGGGCGGTCATTAATGCAGGCGTTCTACTGGACGTG 63

DB 530 GCAAGATATTCGCACTCCGAGAGCGCGGCTTAAATGCAGGCGTTCTACTGGACGT 589

QY 64 CTTTCAGGAGGATATGTTGGGACACATACCGGAGAGATACCGAGTGGTACGATGCC 123

DB 590 CCAGCGGAGGATCTGTTGGATACATCAGAACAGATACCGAGTGGTACGAGGCT 649

QY 124 GGAATCCGCAATATGATATCCCGCGGAGCAAGGGGATGGCGGCGCTTATTCGATG 183

DB 650 GGAATCCGCGCATCTGATATCCCGCGGAGCAAGGGGATGGCGGCGCTTATTCGATG 709

QY 184 GGTACGACCCCTACGACTCTTTGACCTCGTGTAGTACGACCAAGAGGAGCGTAGAG 243

DB 710 GGTACGACCCCTACGACTCTTTGACCTCGTGTAGTACGACCAAGAGGAGCGTAGAG 769

QY 244 ACGCGCTTTGGTCCAGCAGAGCTCGTGAAATGATAAACAACCGCCACGCTTATGGC 303

DB 770 ACTCGCTTTGGTCCAGCAGAGCTTATCAACATGATAAACAACCGCCACGCTTATGGC 829

QY 304 ATGAGGTATACCGATATAGTATCAACACCGCGCGGCGGCTGACCTGGAGTGGAC 363

DB 830 ATGAGGTATAGTATAGTATCAACACCGCGCGGCGGCTGACCTGGAGTGGAC 889

QY 364 CCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGAGTGGCTCGGCTGAATACACG 423

DB 890 CCCTTCGTGGGACTATACCTGGACCGACTTCTCAAGAGTGGCTCGGCTGAATACACG 949

QY 424 GCACTACTCGACTTCAACCGCAACGATCCATGCGGGGATTCGGAACATTTGGA 483

DB 950 GCACTACTCGACTTCAACCGCAACGATCCATGCGGGGATTCGGAACATTTGGA 1009

QY 484 GGTATCCGACATATGCAACAGAGCTGGGACGACTGGCTCTGGCGGAGCGGAC 543

DB 1010 GGTATCCGACATATGCAACAGAGCTGGGACGACTGGCTCTGGCGGAGCGGAC 1069

QY 544 GAGAGTACCGGCAATATTCAGAGGATTCGAGCATGATGCGGCTTCGATACGTC 603

DB 1070 GAGAGTACCGGCAATATTCAGAGGATTCGAGCATGATGCGGCTTCGATACGTC 1129

QY 604 AAGGGCTATGCTCCCTGGGCTCGTCAAGGACTGGCTGAATGTTGGGAGGCTGGCGGTT 663

DB 1130 AAGGGCTACGGAGCGTGGGCTCGTCAAGGACTGGCTGAATGTTGGGAGGCTGGCGGTT 1189

QY 664 GGAGAGTACTGGGACACCAACGCTCGAGCGTGTCTCACTGGGCACTATCGAGCGGTGCC 723

DB 1190 GGTGAGTACTGGGACACCAACGCTGTGATGCACTCTCACTGGGCACTATCGAGCGGTGCC 1249

QY 724 AAGGTCTTTGACTTCCGCTCTACTCAAGATGGATGAGGCTTTGACAAACAAACAT 783

DB 1250 AAGGTCTTCCGCTCTACTCAAAATGGAGAGGCTTTGACAAACAAACAT 1309

QY 784 CGAGCGCTCGTCTCGGCTTCAGACGCGCAGACTGTGTCTCCGCGACCGCTTCAAG 843

DB 1310 CGGCGCTTGGTTGATGCGCTTCAGAAACCGGGGAAACGTCGTCTCTCGGACCGCTTCAAG 1369

QY 844 GCGTAAACCTTTGTAGCAAAACCAACGACCGCATATAATCTGGAACAAGTATCCAGCTAC 903

DB 1370 GCGTAAACCTTTGTAGCAAAACCAACGACCGCATATAATCTGGAACAAGTATCCAGCTAC 1429

QY 904 GCGTAACTCTACCTACGAGGCGCAGCGCAATATTTCTACGCGACTACGAGGAGTGG 963

DB 1430 GCTTTTCATCTCCTACCTACGAGGCGCAGCGCGTCATATTTCTACGCGACTACGAGGAGTGG 1489

QY 964 CTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGAGGAGAAC 1023

DB 1490 CTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGAGGAGAAC 1549

QY 1024 ACCGACATAGTCTACTACGATACGATGAACCTCATCTTCGTCAGGAAACGCTACGCGGAC 1083

DB 1550 ACCGACATAGTCTACTACGATACGATGAACCTCATCTTCGTCAGGAAACGCTACGCGGAC 1609

QY 1084 AAGCGGGGCTTATTAACCTACATCAACCTAGCTCGAGCAAGCGCGGAAAGTGGTTTAT 1143

DB 1610 AAGCGGGGCTTATTAACCTACATCAACCTAGCTCGAGCAAGCGCGGAAAGTGGTTTAT 1669

QY 1144 GTCCGCAAGTTCGCGGCGCTGCATCCACAGTACACCGGCAACCTCGGAGGCTGGGTA 1203

DB 1670 GTCCGCAAGTTCGCGGCGCTGCATCCACAGTACACCGGCAACCTCGGAGGCTGGGTA 1729

QY 1204 GACAAGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCGCTGCC 1263

DB 1730 GACAAGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCGCTGCC 1789

QY 1264 AAGCGGCAAGTTCGCGGCGCTGCATCCACAGTACACCGGCAACCTCGGAGGCTGGGTA 1311

DB 1790 AAGCGGCAAGTTCGCGGCGCTGCATCCACAGTACACCGGCAACCTCGGAGGCTGGGTA 1837

## RESULT 8

AF017454

LOCUS

DEFINITION

AF017454

ACCESSION

AF017454.1

VERSION

GI:2655009

KEYWORDS

ORGANISM

Thermococcus sp. Rt3

Thermococcus sp. Rt3

Archaea; Euryarchaeota; Thermococci; Thermococcales;

Thermococcaceae; Thermococcus.

1 (bases 1 to 2372)

Jones, R.A., Jermin, L.S., Eastal, S., Patel, B.K. and Beacham, I.R.

Amlyase and 16S rRNA genes from a hyperthermophilic archaeobacterium

J. Appl. Microbiol. 86 (1), 93-107 (1999)

99154282

MEDLINE

PUBMED

10030014

REFERENCE

2 (bases 1 to 2372)

Jones, R.A., Patel, B. and Beacham, I.R.

Direct Submission

Submitted (06-AUG-1997) School of Biomedec. and Biomedical Science,

Griffith University, Nathan, Brisbane, Qld. 4111, Australia

Location/Qualifiers

1. .2372

/organism="Thermococcus sp. Rt3"

/mol\_type="genomic DNA"

/strain="Rt3"

/db\_xref="taxon:65421"

gene	372. .1781	
	/gene="amy"	
CDS	372. .1781	
	/gene="amy"	
	/EC_number="3.2.1.1"	
	/codon_start=1	
	/transl_table=11	
	/product="amylase"	
	/protein_id="AAB87860.1"	
	/db_xref="GI:2655010"	
	/translation="MSAKKLLALLFVLAVGVAVIPARVGIAPVSAGATSRPSLEEG	
	GVIMQAFYWDVPAGGIWDTIRSKIPWASAGISALWIPSPKSGMSGRISMGIDPYDF	
	FDLGEYKQGVETRFSGKQLINIMTAHSYGIKVIADIVINHRAGGDLWNPFTNS	
	YTWDFSKVAGTFTANYLDHPNEKCCDEGTFGFPDIAHEKSWDQYWLWASQKSY	
	AAYLRSIGIDAWRFDYKGYCAWYKDLKWLALVAGYVNDTNVDLHLLWYSSGAKV	
	FFFLYKMDFAFNKIPALVSLQNGQVVRDPFKAVTFVANHDTDIIWNKYPAY	
	AFILTYGQPIVFRDYDEEMLNKLRLNLIWHNNLAGGSTSIYVYDNDLIFVRNGY	
	GNKPLITVILNGSKVGRWYVVPKFAAGSIHETGNLGHWDXVYVSGNSWVLEAPA	
	HDPKAGQGYISWNSYCGVG"	
ORIGIN		
Query Match	73.5%; Score 963.6; DB 1; Length 2372;	
Best Local Similarity	84.4%; Pred. No. 3.4e-207;	
Matches 1096; Conservative	0; Mismatches 199; Indels 3; Gaps 1;	
Qy	14	CGAGCTGGAAGGGCGGGTCTAATGCGAGGGCTTCTACTGGAGCGTGCCTTCAGAG 73
Db	487	CGAGCTCGAAGAGGGCGGGTCTAATGCGAGGCTTCTACTGGAGCGTGCCTTCAGAG 546
Qy	74	GAATATGTGGGACACAATACGGCAGAGATACCGAGTGTGATCGATGCCGGGAATCTCCG 133
Db	547	GAATCTGTGGATACATATAGGAGCAAGATACCGGACTGGCCAGCGCTGGAACTCTAG 606
Qy	134	CAATATGGAATCCCGCGAGCAAGGCATGGCGCGCTTATTCGATGGGCTTACGACC 193
Db	607	CTATATGGAATCCCTCCCGAGCAAGGGTATAGTGGTGGCTTATTCATGGGCTTACGACC 666
Qy	194	CCTACGACTCTTTGACCTCGTGAGTACGACGAGGAGGACGATAGAGCGGCTTTG 253
Db	667	CCTACGACTCTTTGACCTCGGCGAGTACTACGAGAGGAGGACGTTGAGACCCGCTTTG 726
Qy	254	GCTCAAGCAGGAGCTCGTGAACATGATAACACCGCCAGCCCTATGCGATGAAGGTAA 313
Db	727	GCTCAAGCAGGAGCTAATTAACATGATAAACAACCGCCACTCTCTATGGAATCAAGGTCA 786
Qy	314	TAGCGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGGAACCCCTTCTGTGA 373
Db	787	TAGCGACATAGTATCAACACCGCGCGCGGTGACCTGGAGTGGAACCCCTTCTCAA 846
Qy	374	ACGACTATACCTGGACCGACTTCTCAAAGGTGGGTGGGTAAATACAGGCCCAACTACC 433
Db	847	ACAGCTACCTGGAGTACTCTCAAGGTGGCTCGGCGAGTACACCGCCAACTACC 906
Qy	434	TCGACTTCCACCGCAACGAGTCCATGCGGCGATTCGGAACATTTGGAGGCTATCCCG 493
Db	907	TCGACTTTCACCCCAACGAGTCAAAATGCTGTGATGAAGGCACCTTCGAGGGTTCGCG 966
Qy	494	ACATATGCCAGCAGAGCTGGGACGATGCTGCTGGCGCAGCCAGCAGGAGCTACG 553
Db	967	ACATAGCCCATGAGAAGAGCTGGGACCGAGTACTGCTCTGGCCAGCCAGAGAGCTATG 1026
Qy	554	CGGCATATCTCAGGAGCATCGGCTCGATGCTGGCGTTCGACTAGTCAAGGGCTATG 613
Db	1027	CTGCCTACCTCAGGAGCATAGGCTCGATGCTGGCGCTTCGACTAGTCAAGGGTTCAG 1086
Qy	614	CTCCCTGGGTCTCAGAGCTGGCTGACTGGTGGGAGGGCTGGCGGGTTGGAGAGTACT 673
Db	1087	GGGCGTGGGTGGTCAAGAGCTGGCTGAAATGGTGG---GGCGTGGCGGTGGTGGAGTACT 1143
Qy	674	GGGACACCAAGCTGACCGCTGTCTCAACTGGGATACTCGAGCGGTGCCAAGGTCTTTG 733
Db	1144	GGGATACAAAGTTGATGCCCTCTTAATCGGCGCTACTCAAGGGGTGCCAAGGTCTTCG 1203

Qy	734	ACTTCGCCCTCTACTACAAGATGGATGAGCGCTTTTGACAAACAAAACATTTCCAGCGCTCG 793
Db	1204	ACTTCGCCCTCTACTACAAGATGGATGAGCGCTTTTGACAAACAAAACATTTCCAGCGCTCG 1263
Qy	794	TCTTCGCCCTTCAGAACCGCCAGACTGTGTGCTCCCGCGACCGGTTCAAGCGCGTAACT 853
Db	1264	TTTCGCCCTTCAGAACCGGTTCAGACAGTGTGTTCGCCGACCCCTTTCAAGGGTGTAACT 1323
Qy	854	TTCTAGCAAAACACGACCGCATATATCTGGAACAAAGTATCCAGCTTACGGTTCATCC 913
Db	1324	TGTCGCCCAACACGACCGCATATCTGGAACAAAGTATCCGGCTTATGCTTCATCC 1383
Qy	914	TCACCTACGAGGCCAGCCGACAAATTTCTACCGGACTACGAGGAGTGGCTCAACAAG 973
Db	1384	TCACCTACGAGGCTCAGCGGTTATATTCTACCGGACTACGAGGAGTGGCTCAACAAG 1443
Qy	974	ATRAAGCTCAAGAACCTCATCTGGATACATGAACTCCCGCGGAGGAACGACGACATAG 1033
Db	1444	ACAGCTCAAGAACCTCATCTGGATACACAACTCTCCGCCGCGGAGCAGAGCATAG 1503
Qy	1034	TTTACTACGATACGATGAACCTATCTTCGTGAGGAACCGGCTACGGGGACAAGCCCGGGC 1093
Db	1504	TTTACTACGATACGAGGAGCTCATCTTCGTGAGGAACCGGCTACGGGAACAAGCCAGGCC 1563
Qy	1094	TTTATACCTACATCAACCTAGGCTCGAGCAAGCCCGAAGGTGGGTTCATGTCGCAAGT 1153
Db	1564	TTTATACCTACATCAACCTCGGCTCCAGCAAGGTGGAAGGTGGGTTCATGTCGCAAGT 1623
Qy	1154	TCGCGGCGGCTGCTACCTCACAGTATACCTGCTGAGGAGCTGGGTGGGTGGGTTCATGTCGCAAGT 1213
Db	1624	TCGCGGCTTCGTGCTATCCAGGATACACGGGCAACCTCGCGGCTGGGTTCACAGTACG 1683
Qy	1214	TTTACTCAAGCGGCTGGGTCTATCTCGAAGCTTCAGCTTACGACCTGCAACGCGGAGT 1273
Db	1684	TTGGCTCAAAACGAGTGGGTCTACCTCGAGGCGCTGCCACGACCTGCGCAAGGGCAGT 1743
Qy	1274	ATGGCTACTCCGTGGAGCTACTCGGGGTGGGTGA 1311
Db	1744	ACGGCTACTCCGTGGAGCTACTCGGGGTGGGTGA 1781
RESULT 9		
AF177906		
LOCUS	AF177906	1480 bp DNA linear BCT 10-SEP-1999
DEFINITION	Pyrococcus woesei alpha amylase (amyA) gene, complete cds.	
ACCESSION	AF177906	
VERSION	AF177906.1	GI:5853153
KEYWORDS		
SOURCE	Pyrococcus woesei	
ORGANISM	Pyrococcus woesei	
	Archaea; Euryarchaeota; Thermococci; Thermococcales;	
	Thermococcaceae; Pyrococcus.	
REFERENCE	1 (bases 1 to 1480)	
AUTHORS	Chunlin Lu, Weizheng, J. and Yunyan, Y.	
TITLE	Cloning and Expression of Alpha Amylase from Pyrococcus woesei	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 1480)	
AUTHORS	Chunlin Lu, Weizheng, J. and Yunyan, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-AUG-1999) New Strategy Biotech Company, 500 Cao Bao	
	Road, Shanghai 200233, P.R.China	
FEATURES	Location/Qualifiers	
source	1. .1480	
	/organism="Pyrococcus woesei"	
	/mol_type="genomic DNA"	
	/strain="DSM5773"	
	/db_xref="taxon:2262"	
gene	98. .1480	
	/gene="amyA"	
CDS	98. .1480	
	/gene="amyA"	
	/SC_number="3.2.1.1"	
	/note="similar to Pyrococcus furiosus alpha amylase"	

```

/codon_start=1
/transl_table=11
/product="alpha amylase"
/protein_id="AADS4338.1"
/db_xref="GI:5853154"
/translation="NNIKKLPTLLTLLFFIVLSPVSAKYLEBEGGVIMQAFYWD
VPEGGIWDHIRSKIPVEYAGISAIWLPPSKMSGGYSDYDIDFGLGEYQKG
VTEFRGSEELVRLIOTAHAYGIKVIADVINHRAGDLWNPVGGDYTWDFSKVA
SKYTYANLDFHNLHCCDECTFGFPICHKEWDOYWLKSNESVAAVLRISGPD
GWRFDYKGYGAVRDMLNHWGNAVGEYWDVNDALLSWAYSGAKVDFPFPIYYKM
DEAFONNNIPALVIALQNGQVVSDFKAVTFVANHDIDIIWNKIPAYAFILYEGQ
PVIFRDFEELNKLINLIWHDELGGSTTVIYDNDDELI FVRNGDSRRPLGITY
INLSPNWGRWYVVPKFAACIHEYTGNLGNWVKRVDSSGVVLEAPHPDPANGYVG
YSVWSYCVGVG"

ORIGIN
Query Match          50.5%; Score 792.8; DB 1; Length 1480;
Best Local Similarity 75.4%; Pred. No. 1.3e-168;
Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

QY 4 GCCAAGTACTCCGAGCTGGAAGGGCGGGGTCAATAATGCAGCGCTTCTACTGGGACGTG 63
DB 173 GCARAATACTTGGAGCTTCGAGGGGAGGAGTATAATGCAAGCATTCATTGGGATGT 232
QY 64 CTTTCAGGAGGAATATGGTGGGACACATACCGCGAGAGATACCGGAGTGGTACGATGCC 123
DB 233 CCAGGGGGAGGAATTTGGTGGGATCATATAAGATCGAAGATTCCTGAATGGTATGAAGCT 292
QY 124 GGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGGCGCTTATTCGATG 183
DB 293 GGAATCTCTGCAATATGGCTACTCCACCAAGCAGGGATGAGTGGAGGATATTCAATG 352
QY 184 GGCTACGACCCCTACGACTTCTTTGACCTCGTGGTACGACGAGAGGAAACGGTAGAG 243
DB 353 GGCTACGATCCCTATGATTACTTTGATCTCGGCGAGTACTACGAGAGGAACTGTAGAG 412
QY 244 ACGCGCTTTGGCTCCAGCAGAGCTGCTGAACNTGATAAACAACCGCCCGCTATGCC 303
DB 413 ACGCGTTTTGGATCAAAAAGAACTAGTGGATTGATACAACTGCCATGCCCTATGGA 472
QY 304 ATGAAGGTAATAGCCGATATAGTCATCAACCCGCGCGGGGTGACCTGGAGTGGAAAC 363
DB 473 ATAAAGGTAATCGCCGATGTAGTTATAAACCACAGGGCTGGTGGTGAACCTAGAAATGGAAC 532
QY 364 CCCTTCGTGAACACTATACCTGGACCGACTTCTCAAAGTCCGCTCGGTTAAATACAG 423
DB 533 CCCTTCGTGGAGATTACACATGACAGACACTTTCTTAAAGTTGCTCAGGGGAAATATACA 592
QY 424 GCCAACTACTCTCGACTTCCACCCGACGAGCTCCATGCGGGGCTTCCGGAACATTGGA 483
DB 593 GCTAACTACTGACTTCCATCCAAACGAGCTTCATTGTTGACGAGGAACCTTTGGA 652
QY 484 GGCTATCCCGACATATGCCACGACAGAGCTGGACCACTATGCTCTGGGCCACGCCAG 543
DB 653 GGATTTCCAGATATATGTATCATCAAAAGAGTGGATCAGTACTGCTATGGAAGAGCAAT 712
QY 544 GAGAGTACGCGGCATATCTCAGGAGCATCGGCATCGCTCGGCTTCGACTACGTC 603
DB 713 GAGAGTTATGCTGCTATTATTAAGAACATAGATTTGATGGTTGAGATTGACATGTT 772
QY 604 AAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGCGGTT 663
DB 773 AAGGGCTATGGAGCTTGGGTTGTCAGAGACTGGCTTAATGGTGGGAGGTTGGGCGATT 832
QY 664 GAGAGTACTGGGACACCAACCTGCGAGCTGTCTCAACTGGGCATCTCGAGCGGTGCC 723
DB 833 GAGAGTACTGGGACACAAATGATGATGCACTACTAAGCTGGGCATATGAGAGTGGTGCA 892
QY 724 AAGGTCTTTGACTTCGCGCTTACTACTAAGAGTGGATGAGCGCTTTTGACAAACAAACATT 783
DB 893 AAGGTCTTTGACTTCGCGCTTACTACTATAAATGGATGAAGCATTTTGACAAATCAACATT 952
QY 784 CCAGCGCTGCTCTCTGCGCTTCAGAAAGCGGCAGACTGTTGCTCTCCGCGACCGGTTCAAG 843

```

```

Db 953 CCAGCATTAAGTCTATGCCCTACAAAACGGACAACCTGTAGTTTCGAGAGATCCATTAAAG 1012
QY 844 GCGGTAACTTTGTAGCAACACGACGATATTAATCTGGAAACAGATATCCAGCCTAC 903
Db 1013 GCAGTAACTTTGCTGTCCTCAATCAATGACAGATATAATTAATGAACAGATATCCAGCAT 1072
QY 904 GCGTTTCATCTCCCTACGAGGCGCCAGCGACAATAATTTCTACCGACTACGAGGAGTGG 963
Db 1073 GCGTTTCATTTGACATATGAGGACAGCCAGTAATATTTCTACAGGACATTTGAGGAATGG 1132
QY 964 CTCAACAGGATTAAGCTCAAGAACCTTCTATCTGATACATGAGAACCTCGCGGAGGAAGC 1023
Db 1133 CTGAACAAGGATTAAGCTTAATTAACCTCATTTGGATCCATGATCATTTGGCAGGAGGAAGC 1192
QY 1024 ACCGACATAGTCTACTACGATAACGATGAATCATCTTCTGTCAGGAACGGCTACCGGGAC 1083
Db 1193 ACAACATTTGCTACTACGACACGATGAGCTCATATTTGTGAGAATGGAGATTCAGA 1252
QY 1084 AAGCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCCCGAAGTGGGTTTAT 1143
Db 1253 AGGCTGGGCTTATAACCTTACATTAACCTTGAAGCCCTAACCTGGGTTGTTAGGTGGTATAC 1312
QY 1144 GTGCGGAAGTTCCGCGCGGCTGCATCCAGAGTATACCTGTAACCTCGAGGCTGGGTA 1203
Db 1313 GTTCCAAAGTTTCAGGGGCTTGTATTTCTGAATACACTTGAACCTAGGAGGATGGGTA 1372
QY 1204 GACAAGTACCTCTACTCAAGCGGCTGGGCTTATCTCGAAGCTCCAGCTTACGACCTGCC 1263
Db 1373 GATAAAGAGCTAGATAGTAGGATGGTATACCTAGAGGACCACTCAGATCCAGCT 1432
QY 1264 ACGGGCAGTATGGCTACTCCGCTGGAGCTACTCGGGGTTGGGCTGA 1311
Db 1433 ACGGCTACTATGGTACTCCGATGGAGTTATTTGGTGGTGA 1480

RESULT 10
AF240464
LOCUS AF240464 1561 bp DNA linear BCT 21-MAR-2000
DEFINITION Pyrococcus woesei alpha-amylase gene, complete cds.
ACCESSION AF240464
VERSION AF240464.1 GI:7271926
KEYWORDS
SOURCE Pyrococcus woesei
ORGANISM Pyrococcus woesei
Archaea; Euryarchaeota; Thermococci; Thermococcales;
Thermococcaceae; Pyrococcus.
REFERENCE
1 (bases 1 to 1561)
Wang, C. and Lo, H.
Cloning of Pow alpha-amylase gene
Unpublished
2 (bases 1 to 1561)
Wang, C. and Lo, H.
Direct Submission
Submitted (29-FEB-2000) Phamology, Viogenic, P. O. Box 70472,
Sunnyvale, CA 94086, USA
FEATURES
Location/Qualifiers
1..1561
/organism="Pyrococcus woesei"
/mol_type="genomic DNA"
/db_xref="taxon:2262"
/notes="DSM 3773"
137..1519
/notes="pow"
/codon_start=1
/transl_table=11
/product="alpha-amylase"
/protein_id="AAF4453.1"
/db_xref="GI:7271927"

```



Qy	4	GCCAGCTACTCGAGCTGGAAAGGGCGGGTCATATATGACGGCGTTCTACTCGGACGTG	63
Db	212	GCAAAATACTTGGAGCTTGAAGAGGGAGGAGTTATAATGCAAGCATTTCTATTGGGATGTT	271
Qy	64	CTTTCAGAGGAATATGCTGGGACACAATACGGCAGAGAATACCGGAGTGGTACGATGCC	123
Db	272	CCACGGGGAGGAATTTGGTGGGATCATATAAGATCGAAGATTCCTGAATGTTATGAAGCT	331
Qy	124	GGAACTCTCCGAATATGGATTTCCCCCGGAGACGAGGCGATGGCGGGCCCTATTGCGATG	183
Db	332	GGAACTCTCTGCAATATGGCTACCTCCACCAAGCAAGGGGATGATGTGGAGATATTTCAATG	391
Qy	184	GGCTACGACCCCTACGACTTCTTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTAGAG	243
Db	392	GGCTACGATCCCTATGATTATTTTGATCTCGGCGAGTACTCCNAGAGGGNACTGTAGAG	451
Qy	244	ACGGCTTTGGCTCCAAAGAGGAGCTCGTGAAATATGATAAACACCGGCCACGGCTATGGC	303
Db	452	ACCGTTTTTGGATCAAAAGAAAGAACTAGTGAGATTTGATACAAATGCCCCATGCTATGGA	511
Qy	304	ATGAAGGTAAATAGCCGATATAGTTCATCAACACACCGCGCGGGGTGACCTTGGAGTGGAAAC	363
Db	512	ATAAAGTAACTGCCGATGTAGTTATAAACCAAGGGCTGGTGGTGACTTAGAATGGAC	571
Qy	364	CCCTTCGTGAACGACATATACCTGGACCGACTTCTCAAAGTTCGGGTCCGGTAAATACACG	423
Db	572	CCCTTCGTGGAGATTACATGACGACAGACTTTTCTAAAGTTGGCTCAGGGAATAATACA	631
Qy	424	GCCAATACCTCGACTTCCACCGGAAACGAGCTCCATGCGGCGCATTCGGGAAACATTTTGG	483
Db	632	GCTTAATCTCTGGACTTCCATCCAAACGAGCTTCATTGTTGTGACGAAGGAAACCTTTTGG	691
Qy	484	GGCTATCCGACATATGCCACGACAGAGCTGGGACCAAGTACTGGCTCTCGGCCAGCCAG	543
Db	692	GGATTTCCAGATATGTGTCATCAAAAGAGTGGGATCAGTACTGGCTATGGAAGAGCAAT	751
Qy	544	GAGAGCTACGGGATATCTCAGAGAGCATCGGCATCGATGCCTGGCGCTTCGACTAGCTC	603
Db	752	GAGAGTTATGCTGCTTATTTAAAGAGCATAGGATTTGATGGTTGGAGATTGACTATGTT	811
Qy	604	AAGGGCTATGCTCCCTGGGTGCTCAAGGAGCTGGCTGAACCTGGTGGGAGAGCGTGGCGGGT	663
Db	812	AAGGGCTATGAGCTTGGGTGTTCCAGAGACTGGCTTAATGGTGGGAGGTTGGGCGATT	871
Qy	664	GGAGAGTACTGGGACACCAAGTCGACGCTGTTCTTCACTGGGCGATACTCGAGCGGTGCC	723
Db	872	GGAGAGTACTGGGACACAAATGTAGATGCACTACTAAGCTGGGCGATATGAGAGTGGTGA	931
Qy	724	AAGGTCTTTGACTTTCGCCCTCTACTACAAGATGATGAGGCGCTTTGACAAACAAAACATT	783
Db	932	AAGGTCTTTGACTTTCGCCCTCTACTATAAATGGAATGCAATTTGACAAATACCAATT	991
Qy	784	CCAGCGCTCGTCTCTGCCCTTGAAGACGGCGCAGACTGTTGTCTCCCGCGACCCGCTCAAG	843
Db	992	CCAGCATTAGTCTATGCCCCCTACAAACGGGACAACTGTAGTTTCGAGAGATGCCATTTAAG	1051
Qy	844	GCGGTAACCTTTGTAGCAAAACCGACACCGATATAATCTTGGAAACAAGTATCCAGCGCTAC	903
Db	1052	GCAGTAACTTTCGTTGCCAATCATGACACAGATATAATATGGAACAAGTATCCAGCATAT	1111
Qy	904	CGCTTCACTCTACCTACGAGGGCCAGCCGACAAATATTTCTACCGCGACTACGAGGAGTGG	963

Query Match 60.5%; Score 792.8; DB 1; Length 1740;  
Best Local Similarity 75.4%; Pred. No. 1.3e-168;  
Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

QY 4 GCCAAGTACTCCGAGCTGGAAGGGCGGGTGCATTAATGCGAGCTTCTTACTGGGACGTG 63  
DB 304 GCAAAATATCTGGAGCTTGAAGGGGAGGATTAATATGCAAGCAATCTTATTTGGATGTT 363  
QY 64 CTTTCAGGAGGATATGTTGGGAGGACATACCGGAGAGATACCGGAGTGTACGATGCC 123  
DB 364 CCAGGGGGAGGAAATTTGGTGGGATCATATAAGATCGAAGATTTCTGAAATGGTATGAAGCT 423  
QY 124 GGAATCTCCGCAATATGGATTTCCCGGGAGCAAGGGCATGGCGGCGCTTATTCGATG 183  
DB 424 GGAATCTCTGCAATATGGCTACTCCACCAAGAGGGATGAGTGGAGGATATTTCAATG 483  
QY 184 GGTACGACCCCTACGACTTTCTTGAACCTTCGTTAGTACGACCAAGAGGAAACGGTAGAG 243  
DB 484 GGTACGATCCCTATGATTACTTTGATCTCGGCGAGTACTACCAAGAGGAACTGTAGAG 543  
QY 244 ACGCGTTTGGCTCCAGCAGAGCTCGTGAACATGAATAACACCGCCCGCTTATCGC 303  
DB 544 ACGGTTTTGGATCAAAAGAGAACTAGTGAATGATCAAACTGCCATGCCCTATGGA 603  
QY 304 ATGAAGGTAATAGCCGATATAGTCTATCAACACCGCGCGCGGTGACCTGGAGTGAAC 363  
DB 604 ATAAAGGTAATCGCCGATAGTTTATATAACACAGCGGTGGTGGTGAACCTAGATGGAAC 663  
QY 364 CCCTTCGTGAACGACTATACCTTGGACCGACTTCTCAAGGTCGCTCGGTGAATACAG 423  
DB 664 CCCTTCGTGGAGATTACACATGGAAGCACTTTCTTAAAGTTGSCCTCAGGGAATATACA 723  
QY 424 GCCAATCTACTCGACTTCCACCGCAACGAGCTCCATGCGGGCGATTCGGGAACATTTGGA 483  
DB 724 GCTAATCTATCTGACTTCCATCCAAACGAGCTTCATTTGTTGAGCAAGAACTTTGGA 783  
QY 484 GGTATCCGCAATATGCGACAGACAGAGCTGGGACAGTACTGCTCTGGGCGAGCCAG 543  
DB 784 GGAATTTCCAGATATATGTATCAAAAGTGGGATCAGTACTGGCTATGGAAGCAAT 843  
QY 544 GAGAGTACGCGCATATCTCAGGAGCATCGGCATCGATCGCTGGCTTCGACTACGTC 603  
DB 844 GAGAGTATGCTGCTTATTTAAGAGCATAGGATTTGATGTTGAGATTTGACTATGTT 903  
QY 604 AAGGGTATGCTCCCTGGGTCGTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGGTT 663  
DB 904 AAGGGTATGGAGCTTGGGTTGTCAGAGACTGGCTTAATTTGGTGGGAGGTTGGGCACTT 963  
QY 664 GGAGAGTACGGGACACCAACGTCGACGCTGTTCTCACTGGGCATCTCGAGCGGTGCC 723  
DB 964 GGAGAGTACGGGACACAAATGTAGATGCACACTAAGCTGGGCATATAGAGTGGTGCA 1023  
QY 724 AAGGTCTTTGACTTCGCCCTCTACTACAGATGGATGAGCGCTTTGACAAACAAACATTT 783  
DB 1024 AAGGTCTTTGACTTCGCCCTCTACTATAAATGGATGAAGCAATTTGACAAATCAACATTT 1083  
QY 784 CAGCGCTGCTCTGCTCCCTTCAGAAACGCGAGACTGTTGTTCTCCGGGACCCGTTCAAG 843  
DB 1084 CCAGCAATATAGTATGCTCCCTCAAAACCGGACAACTGTAGTTTCAGAGAGATCCATTTAAG 1143  
QY 844 GCGGTAACTTTGTAGCAAAACCAACGACCGATATAAATCTGGAAACAGATATCCAGCTTAC 903  
DB 1144 GAGTAACCTTTGTTGGCAATCATGACACAGATATAAATGAAACAGATATCCAGCATAT 1203  
QY 904 GGTTCATCTCTACCTACGAGGGCCAGCCGCAAAATATTTCTACCGGCACTACGAGGATGG 963  
DB 1204 GCGTTTATATTTGACATATAGGGGACAGCCAGTAATATTTCTACAGGCACTTTGAGGAATGG 1263  
QY 964 CTCACAGAGGATAAGCTCAAGAACTCATCTGGATACATGAACTCGCGGAGGAAAGC 1023  
DB 1264 CTGAACAGGATAGGTAATTAACCTCATTTGGATCCATGATCATTTTGGCAGGAGGAAGC 1323

QY 1024 ACCGACATAGTCTACTACGATAACGATGAACCTCATCTTCGTGAGGAACGGCTACGGGAC 1083  
DB 1324 ACAACATATGTTCTACTACGACAACGATGAGCTCATATTTGTGGAATATGGAGTTCTAGA 1383  
QY 1084 AAGCCGGGGTATTAACTACATCAACCTAGGCTCGAGCAAGCCGGAAGGTTGGTTTAT 1143  
DB 1384 AGGCTGGGCTTATACTTACATTAACCTTGAGCCCTAACTGGGTTGTTAGTGGGTATAC 1443  
QY 1144 GTCCGGAAGTTCCGGGCGGTGCTATCCACGAGTATACCTGTAACCTCGAGGCTGGGTA 1203  
DB 1444 GTTCCAAAGTTTCGAGGGGCTTGTATTCATGAATACACTGGAACCTAGAGATGGGTA 1503  
QY 1204 GACAAGTACCTCTACTCAAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
DB 1504 GATAAAAGAGTAGATAGTAGCGATGGTATACCTAGAGGACCACTCACGATCCAGCT 1563  
QY 1264 ACGGGCAGTATGGCTACTCCGTGCTGGAGTACTGCGGGTGGGCTGA 1311  
DB 1564 AACGGCTACTATGGGTACTCCGTATGGAGTATTGTTGGTGGGTA 1611

RESULT 12  
E09372  
LOCUS E09372 2134 bp DNA linear PAT 29-SEP-1997  
DEFINITION Heat-resistant alpha-amylase gene.  
ACCESSION E09372  
VERSION E09372.1 GI:22025999  
KEYWORDS JP 1995143880-A/1.  
SOURCE Pyrococcus furiosus  
ORGANISM Pyrococcus furiosus  
Archaea; Euryarchaeota; Thermococci; Thermococcales;  
Thermococcaceae; Pyrococcus.  
REFERENCE 1 (bases 1 to 2134)  
AUTHORS Taguchi,Y., Nishikawa,M., Koyama,N., Yamamoto,K., Asada,K. and Katou,I.  
TITLE ALPHA-AMYLASE GENE HAVING ULTRA-HIGH HEAT-RESISTANCE  
JOURNAL Patent: JP 1995143880-A 1 06-JUN-1995;  
TAKARA SHUZO CO LTD  
COMMENT OS Pyrococcus furiosus  
PN JP 1995143880-A/1  
PD 06-JUN-1995  
PF 18-AUG-1994 JP 1994215253  
PI 01-OCT-1993 JP 93P 267777  
PI TAGUCHI YUKI, NISHIKAWA MIO, KOYAMA NOBUHIITO, PI YAMAMOTO KATSUHIKO,  
PI ASADA KIYOZOU, KATOU IKUNOSHIN  
PC C12N15/09,C12N9/28,C12N15/09,C12R1:01,(C12N9/28,C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
FH Key Location/Qualifiers  
FH FT source 1. .2134 /organism='Pyrococcus furiosus' FT promoter  
FT 337. .373  
FT RBS 380. .388  
FT CDS 395. .1777 /product='heat-resistant alpha-amylase' FT  
FT precursor,  
FT sig\_peptide 395. .469  
FT mat\_peptide 470. .1774 /product='heat-resistant alpha-amylase'.  
FEATURES  
source 1. .2134  
Location/Qualifiers  
/organism='Pyrococcus furiosus'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:2261'

ORIGIN  
Query Match 60.5%; Score 792.8; DB 6; Length 2134;  
Best Local Similarity 75.4%; Pred. No. 1.3e-168;  
Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

QY 4 GCCAAGTACTCCGAGCTGGAAGGGCGGGTGCATTAATGCGAGCTTCTTACTGGGACGTG 63

Db 470 GCAAAATATTGGAGCTTGAAGAGGAGGAGTTAATGCAAGCAATCTATTGGGATGTT 529  
QY 64 CTTTCAGAGGAGATATGTTGGGACATACGCGAGAGATACCGGAGTGTATGATGCC 123  
Db 530 CCAGGGGAGGAAATTTGGTGGGATCATATAGATCGAAGATTCCTGAATGTATGAAGCT 589  
QY 124 GGAATCTCGCAATATGATTTCCCGCGGAGCAAGGCGATGGCGGCGCTATTTCGATG 183  
Db 590 GGAATCTCGCAATATGATTTCCCGCGGAGCAAGGCGATGGCGGCGCTATTTCGATG 649  
QY 184 GGTACGACCCCTACGACTTTCTTGACCTCGGTGATGACGACCAAGAGGAAACGGTAG 243  
Db 650 GGTACGATCCCTATGATTTCTTGACCTCGGTGATGACGACCAAGAGGAAACGGTAG 709  
QY 244 AGCGCTTTGGCTCAAGCAGGAGCTCGTGAACATGATATAACACCGCCACGCTATGGC 303  
Db 710 AGCGCTTTGGATCAAAAGAGAGACTAGTGAATGATACAACTGCCCTATGGA 769  
QY 304 ATGAAGGTAATAGCGATATAGTCAACACCGCGCGCGGCGGAGCTCGGATGGAAC 363  
Db 770 ATAAAGGTAATGCGCGATGATGATATAAACACGAGGCTGGTGGTGAACCTAGAATGGAAC 829  
QY 364 CCCTTCGTGAACGACTATACCTGGAACGACTTCTCAAGGTCGCGTGGGTAAATACAG 423  
Db 830 CCCTTCGTGGAGATACATGGAACGACTTCTCAAGGTCGCGTGGGTAAATACAG 889  
QY 424 GCCAACTACCTCGACTTCCACCGGAGAGCTCCATCGCGGCGATTCGGAACATTTGGA 483  
Db 890 GCTAACTATCTGGACTTCCATCCAAACGAGCTTCTTAAAGTGGCTCAGGGAATATACA 949  
QY 484 GGTATCCCGACATATGCAACGAGAGCTGGGACGATGCTGCTGGCGCAGCAG 543  
Db 950 GGAATTCAGATATATGATCATCAAAAGAGTGGGATCAGTACTGGCTATGGAAGAGCAAT 1009  
QY 544 GAGAGCTACGGGCGATATCTCAGAGAGCTCGGATCGATCGCTGGCGTTCGACTACGTC 603  
Db 1010 GAGAGTTATGCTGCTTATTAAGAGCATAGATTTGATGTTGGAGTTTGACTATGTT 1069  
QY 604 AAGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGGCT 663  
Db 1070 AAGGCTATGAGCTTGGGTTGTGAGAGACTGGCTTAAATGGTGGGAGGTTGGGCGATT 1129  
QY 664 GGAGAGTACTGGGACCAACGTCGACGCTGCTTCACTGAGGATGAGGCTTTGACAAACAAACATT 783  
Db 1130 GGAGAGTACTGGGACCAACGTCGACGCTGCTTCACTGAGGATGAGGCTTTGACAAACAAACATT 1249  
QY 784 CCAGGCTCGTCTGCTGCTTCAGAGCGGAGCTGTTGCTCCCGGACCCGTTCAAG 843  
Db 1250 CCAGCATGATGCTATGCGCTTCAAAACGAGCAAACTGTAGTTTCGAGAGATCCATTTAAG 1309  
QY 844 GCGTAACTTTGTAGCAAAACACGACCGATATAATCTGGAACAAGTATCCAGGCTAC 903  
Db 1310 GCAGTAACCTTTGTTGCCATCATGACACAGATATAATGGAACAGTATCCAGCATAT 1369  
QY 904 GGTTCATCTCCTACCTAGAGGCGAGCGGACATATTTCTACCGGACTACGAGAGTGG 963  
Db 1370 GGTTCATATTTGACATATGAGGAGCAGCAGTAATATTTCTACAGGAGCTTTGAGGAATGG 1429  
QY 964 CTCACAGAGGTAAGCTCAAGAACTCATCTGGATACATGAGAACCTCGCGGAGGAGG 1023  
Db 1430 CTGAACAGAGTAAGCTAATTAACCTCATTTGGATCCATGATCATTTGGCAGGAGGAGC 1489  
QY 1024 ACCGACATAGTCTACTAGATAAGGATGAATCTATCTTGTGAGAAACGGCTACGGGAC 1083  
Db 1490 ACAACATTTGCTACTAGCAACAGGATGAGCTCATATTTGTGAGAAATGGAGATTTCTAGA 1549  
QY 1084 AAGCGGGGCTTATTAACCTACATCACTAGCTCGACCAAGCGCGGAGGCTTTAT 1143

Db 1550 AGGCTGGGCTTATACTTACATTAACCTTGAAGCCCTAACTGGGTTGGTAGGTGGGTATAC 1609  
QY 1144 GTCCGGAAGTTCGCGGCGGCTGATCCAGAGTACTGCTAACTCCGAGGCTGGGTA 1203  
Db 1610 GTTCCAAAGTTTCGAGGCGCTTGTATTCATGATACCTAGGAACTAGAGATGGGTA 1669  
QY 1204 GACAAGTACGCTTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
Db 1670 GATAAAGAGTAGATAGTAGCGGATGGGTATACCTAGAGGCAACCTCAGGATCAGCT 1729  
QY 1264 AAGGCGAGTATGCTACTCCGCTGCTGAGCTACTCGGGTGGGCTGA 1311  
Db 1730 AACGGCTACTATGGGTACTCCGCTATGGATTTATGTTGGGTGA 1777

## RESULT 13

AF001268

LOCUS

Pyrococcus furiosus alpha amylase (amyA) gene, complete cds.

AF001268

VERSION

AF001268.1

KEYWORDS

SOURCE

ORGANISM

Pyrococcus furiosus

Achaeta, Euryarchaeota; Thermococci; Thermococcales;

Thermococcaceae; Pyrococcus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 2627

/organism="Pyrococcus furiosus"

/mol\_type="genomic DNA"

/strain="DSM 3638"

/db\_xref="taxon:2261"

676.2058

/gene="amyA"

/genes="amyA"

/EC\_number="3.2.1.1"

/note="alpha-1,4-glucan 4-glucanohydrolase; acidic extracellular enzyme"

/codon\_start=1

/transl\_table=11

/product="alpha amylase"

/protein\_id="AAC45663.1"

/db\_xref="GI:2183106"

/translation="MNKKLTLLTLLFFVLASPVSAKYLEEELGGMVQAFVWD

VPGGFWMDHRSKIPETAGISALVPPSGSGSGSMGTDYDFDLGRIYKQK

TVETFRGSEBELVRIQTAYHGIKVIADVINHRAAGDLEWNPFGVDTWTFDSKVA

SKRYTANYLDLPHNELHCCDEGTGFGFDICHHKENDQYWLKSNESYAAVLRISGFD

GKRFYVKGYGAVVVDMLNMWGWAGVYDNTVDALLSWAYESGAKVDFPLPYKGM

DFAFDNNIPALVALONGQTVVSRDPKAVTFVANHDTDIWNKYPAVAFILTYEQQ

PIVFYDFREELNKKLNLWIHDLAGGSTTVVYDNDLIVRNGDSRSPGLTY

INLSNFWGRWVYVPKFPAGACIHEITGNLGNVDRVSSGHWYLEAPHPDPANGYYG

ISVWSYCGVG"

## ORIGIN

Query Match 60.5%; Score 792.8; DB 1; Length 2627;

Best Local Similarity 75.4%; Pred. No. 1.4e-168;

Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

QY 4 GCCAAGTACTCCGAGCTGGAAAGGGCGGGTTCATTAATGACGCGTTCCTACTGGGACGTG 63

Db 751 GCAAAATACCTTGAGCTTGAAGGGAGGATTAATAATGCAAGCATTTCTATTGGGATGTT 810  
QY 64 CTTTACGAGGATATGTTGGGACACATACGGCAGAGATACCGAGTGGTACGATGCC 123  
Db 811 CCAGGGGAGGAAATTTGGTGGGATCATATAAGATCGAAGATTCTTGAATGGTATGAAGCT 870  
QY 124 GGAATCTCCGCAATATGGAATTCCTCCCGGAGCAAGGGGATCGGGGGGCGCTTATTCGATG 183  
Db 871 GGAATCTCTGCAATATGCTACTCTCCACCAAGCAGGGGATGAGTGGAGGATATTCATG 930  
QY 184 GGTACGACCCCTACGATCTCTTTGACCTCGTGGTACGACAGAGGGAACGGTAGAG 243  
Db 931 GGTACGATCCCTATGATTTGATCTCTCGGAGTACTACCAAGGGAACCTGTAGAG 990  
QY 244 AGCGCTTTGGCTCCAGCAGAGCTCGTGAACATGATTAACACCGCCACCGCTTATGCG 303  
Db 991 AGCGCTTTGGATCAAAAGAGAACTAGTGAGATTGATACAACTGCCATGCCATATGGA 1050  
QY 304 ATGAAGGTAATAGCCGATATAGTCATCAACCCGCGCGGGGTGACCTGGAGTGGAAAC 363  
Db 1051 ATAAAGGTAATCGCCGATGTAGTTATAAACCACAGGGCTGGTGTGACCTAGATGGAAC 1110  
QY 364 CCCTTCGTGAACATATACCTGGACCGACTTCTCAAGGTGCGCTCGGTAATACAG 423  
Db 1111 CCCTTCGTGGAGATTACATAGGACAGACTTTCTAAAGTTGCCCTCAGGGAATATACA 1170  
QY 424 GCCAATCTACCTGACATTCACCCGACGAGCTCCATGCGGGGATTCGCGAATTTGGA 483  
Db 1171 GCTAATCTGACATTCATCCAAAGAGCTTCATTTGTTGACGAAGAACCTTTGGA 1230  
QY 484 GGTATCTCCGACATATGCGACGACAGAGCTGGGACCACTAGTGGCTCTGGCCAGCCAG 543  
Db 1231 GGAATTCAGATATATGTCATCAAAAGAGTGGGATCAGTCTGGCTATGGAAGCAAT 1290  
QY 544 GAGAGCTACCGGCATATCTCAGGACATCGGCATCGATGCTGGCGCTTCGACTACGTC 603  
Db 1291 GAGAGTATGCTGCTATTATTAAGACATAGGATTTGATGTTGGAGATTGACTATGTT 1350  
QY 604 AAGGGCTATGCTCCCTGGTGTCAAGGACTGCTGAACTGTGGGAGCTGGGCGGTT 663  
Db 1351 AAGGGCTATGGAGCTTGGTGTGACAGACTGCTGTAATTTGGTGGGAGGTTGGCGAGTT 1410  
QY 664 GAGAGTACTGGGACACCAAGCTGAGCGTGTCTCAACTGGGCATCTCGAGCGGTGCG 723  
Db 1411 GAGAGTACTGGGACACCAATGATGATGCACTAAGCTGGGCATATGAGAGTGGTGA 1470  
QY 724 AAGGTCTTTGACTTCGCGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAACATT 783  
Db 1471 AAGGTCTTTGACTTCGCGCTCTACTATAAATGGATGAAGCATTTGACAAATACCAACT 1530  
QY 784 CAGCGCTGCTGCTGCGCTTGAAGCGGACAGCTGTTCTCCCGCGACCGCTTCAAG 843  
Db 1531 CCAGCATTAGTCTATGCCCTTACAAAACGGAACCTGTAGTTTCGAGAGATCCATTTAG 1590  
QY 844 GCGTAACCTTTGTAGCAACACGACACCGATATATCTGGAACAAAGTATCCAGCTAC 903  
Db 1591 GCAGTAACCTTTGTTGCCATCATGACACAGATATATATATGAAACAGTATCCAGCAT 1650  
QY 904 GCGTTCATCTCTACCTACGAGGCGCAGCGACAAATATCTACCGCATTCAGGAGTGG 963  
Db 1651 GCGTTCATATGACATATGAGGACAGCCAGTAATATCTCAGGAGCTTTGAGGAATGG 1710  
QY 964 CTCACAGAGATAGCTCAAGAACCTCATCTGATATACATGAGAACCTCGCCGAGGAGC 1023  
Db 1711 CTGAAACAGATTAAGCTAATTAACCTCATTTGATTCATGATCATTTTGCAGAGGAAGC 1770  
QY 1024 ACCGACATAGTCTACTACGATAACGATGAACTCATCTTCTGTCAGGAACGGCTACGGGAC 1083  
Db 1771 ACACAATTTGCTACTACGACCAAGTACGCTCATATTTTGTAGAAATGAGATTTCTAGA 1830  
QY 1084 AAGCGGGCTTATACCTACATCACTACCTAGGCTCGAGCAAGCGCCGGAAGGTTGTTAT 1143

Db 1831 AGGCTGGGCTTATACTTACATTAACCTTGAAGCCCTAACCTGGGTTGGTGGGTATATAC 1890  
QY 1144 GTGCCGAAGTTCCGGCGGCTGCATCCAGAGTATACTGGTAACCTCGAGGCTGGGTA 1203  
Db 1891 GTTCCAAAGTTTGGAGGGGCTTGTATTATCATATACACTGGAACCTAGGAGATGGTA 1950  
QY 1204 GACAAGTACCTCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1263  
Db 1951 GATAAAAGATAGATAGTAGGATGGGTATACCTAGAGGACCACTCAGCATCCAGCT 2010  
QY 1264 AAGGGGAGTATGGCTACTCCGTGTGGAGCTACTGCGGGGTGGGCTGA 1311  
Db 2011 AAGGGTACTATGGTACTCCGTATGAGGATTATTGGTGTGGTGA 2058  
RESULT 14  
BD012151/c 3519 bp DNA linear PAT 02-AUG-2002  
LOCUS Polypeptide.  
DEFINITION  
ACCESSION BD012151  
VERSION BD012151.1 GI:22092340  
KEYWORDS WO 0109348-A/4.  
SOURCE Pyrococcus furiosus  
ORGANISM Pyrococcus furiosus  
Archaea; Euryarchaeota; Thermococci; Thermococcales;  
Thermococcaceae; Pyrococcus.  
REFERENCE 1 (bases 1 to 3519)  
AUTHORS Koyama,N., Okui,T., Takakura,H., Asada,K. and Kato,I.  
TITLE Polypeptide  
JOURNAL Patent: WO 0109348-A 4 08-FEB-2001;  
TAKARA SHUZO CO LTD,NOBUTO KOYAMA,TOSHITAKE OKUI,HIKARU TAKAKURA,  
KIYOZO ASADA,IKUNOSHIN KATO  
OS Pyrococcus furiosus  
PN WO 0109348-A/4  
PD 08-FEB-2001  
PF 26-JUL-2000 WO 2000JP004956  
PR 02-AUG-1999 JP 99P 218778  
PI NOBUTO KOYAMA,TOSHITAKE OKUI,HIKARU TAKAKURA,KIYOZO ASADA,PI  
IKUNOSHIN KATO  
PC C12N15/56,C12N9/26,C12P19/14,C12N1/15,C12N1/19,C12N1/21,PC  
C12N5/00  
CC  
FH Key Location/Qualifiers.  
1..3519  
/organism="Pyrococcus furiosus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2261"  
FEATURES  
source  
Query Match 60.5%; Score 792.8; DB 6; Length 3519;  
Best Local Similarity 75.4%; Pred. No. 1.4e-168;  
Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;  
ORIGIN  
4 GCCAAGTACTCCGAGCTGGAAAAGGGCGGGGTCTAATGCAAGCGTTCCTATTGGGATGTT 63  
3427 GCAAATATCTGGAGCTTGAAGAGGGAGGATTATATGCAAGCATTTCTATTGGGATGTT 3368  
64 CTTTCAGGAGGATATGGTGGGACACATACGGCAGAGATACCGGAGTGGTACGATGCC 123  
3367 CCAGGGGAGGAAATTTGGTGGGATCATATAGATTCGAAGATTCTGATGGTATGAGCT 3308  
124 GGAATCTCCGCAATATGGAATTCCTCCCGGCGAGCAAGGGCATGGCGGCGCTTATTCGATG 183  
3307 GGAATCTCTGCAATATGCTACTCTCCACCAAGCAAGGGGATGAGTGGAGGATATTCATG 3248  
184 GGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGGAACGGTAGAG 243  
3247 GGCTACGATCCCTATGATTTACTTTGATCTCGGCGAGTACTACCAAGAGGAACTGTAGAG 3188  
244 ACGGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACCCGCCACGCTATGCG 303  
3187 ACGGGTTTGGATCAAAAGGAAGAACTAGTGAGATTGATACAAACTGCCCATGCTATGGA 3128

QY	304	ATGAGGTAATAGCGATATAGTATCAACCAACCGCGCGGTGACCTGGAGTGAAC	363
Db	3127	ATAAAGGTAATCGCGATGTAGTTATAAACACACAGCGGTGTGTGACCTAGATGAAC	3068
QY	364	CCCTTCGCGAAGCATATACCTTGGACCGACTTCTCAAGAGTCCGGTCGGTAAATACAG	423
Db	3067	CCCTTCGTTGGAGATACACATGGACAGACTTTCTAAAGTTGCCTCAGGGAATATACA	3008
QY	424	GCCAACTACCTCGACTTCCACCCGACGAGCTCCATGCGGGCGATTCGCGAACATTTGGA	483
Db	3007	GCTAACTATCTGGACTTCCATCCAAACGAGCTTCATTGTTGACGAAGAACCTTTGGA	2948
QY	484	GGCTATCCGACATATGCCACGACGAGCTGGACAGTACTGGCTCTGGCCAGCCAG	543
Db	2947	GGATTTCAGATATATGATCATCACAAAGAGTGGGATCAGTACTGGCTATGGAAGAGCAAT	2888
QY	544	GAGAGCTACGCGGCATATCTCAGGAGCATCGGCATCGATGCTGGCGTTCGACTACGTC	603
Db	2887	GAGAGTATGCTGCTTATTAAAGACATAGGATTTGATGTTGGAGATTTGACTATGTT	2828
QY	604	AAGGCTATGCTCCCTCGGTGTCAAGGACTGCTGAACCTGGTGGGAGGCTGGCGGTT	663
Db	2827	AAGGCTATGGAGCTTGGGTTGTCAAGAGCTGGCTTAATTGGTGGGAGGTTGGGCAGTT	2768
QY	664	GGAGAGTACTGGACACCAACGTCACGCTGTCTCAACTGGGCATCTCGAGCGGTGCC	723
Db	2767	GGAGAGTACTGGACACCAAAATGTAGTGCATCTAAGCTGGGCATATGAGAGTGTGCA	2708
QY	724	AAGGCTTTGACTTCGCCCTCTACTACAAGATGGATGAGGCTTTGACAAACAAACATT	783
Db	2707	AAGGCTTTGACTTCGCCCTCTACTATAAATGGATGAACATTGACAAATACAAACATT	2648
QY	784	CCAGCGCTCGTCTCGCCCTTCAGAACCGCCAGACTGTCTCCCGGACCGCTTCAG	843
Db	2647	CCAGCATTAGTCTATGCCCTCAAAACGGACAACTAGTATTGAGAGATCCATTTAAG	2588
QY	844	GCGTAACTTTGTAGCAAAACACACACCGCATATAATCTGGAAACAGATCCAGCCTAC	903
Db	2587	GCAGTAACTTTGTTGCCAATCATGACACAGATATATATGGAACAGTATCCAGCATAT	2528
QY	904	GGTTTCATCTCCTACCTACGAGGCGCAGCGCAATATTTCTACCGGACTACGAGAGTGG	963
Db	2527	GGTTTCATATTTGACATATGAGGCGACGCGCAGTAATATTTCTACGCGGACTTTGAGGAATGG	2468
QY	964	CTCAACAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGACCTCGCGGAGGAGGC	1023
Db	2467	CTGAACAAGGATAAGCTCAATTAACCTCATTTGGATCCATGATCATTTGGCAGGAGGAGC	2408
QY	1024	ACCGACATAGTCTACTACGATAACGATGAATCATCTTCGTCAGGAACGGCTACGCGGAC	1083
Db	2407	ACAACAAATGTCTACTACGACACCGATGAGCTCATATTTGTGAGAAATGGAGATTTCTAGA	2348
QY	1084	RAGCCGGGCTTATAACCTACATCAACCTAGCTCGACAGGCGCGGAGGTGGTTTAT	1143
Db	2347	AGGCTTGGGCTTATAACCTTACATTAACCTTAGCCCTAATCTGGTGGTGGTGGTATAC	2288
QY	1144	GTGCCGAAGTTTCGCGGGCGCTGCATCCACGAGTATCTGTTAACTCGGAGGCTGGGTA	1203
Db	2287	GTTCCTCAAGTTTCGAGGGCTTGTATTCTATGATACATCTGGAACCTAGGAGATGGTA	2228
QY	1204	GACAAGTACGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC	1263
Db	2227	GATAAAGAGTAGTAGTAGCGGATGGGTATACCTTAGAGGACCACTCAGCATCCAGCT	2168
QY	1264	AACGGGCAAGTATGGCTACTCCGTGTGGAGCTACTCGCGGGTGGGCTGA	1311
Db	2167	RACGCTACTATGGGTACTCCGTATGGAGTTATGTGGTGTGGTGA	2120

RESULT 15  
AE010170  
LOCUS

DEFINITION	Pyrococcus furiosus DSM 3638, section 45 of 173 of the complete genome.
ACCESSION	AE010170 AE009950
VERSION	AE010170.1 GI:18892445
KEYWORDS	
SOURCE	Pyrococcus furiosus DSM 3638
ORGANISM	Pyrococcus furiosus DSM 3638 Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae; Pyrococcus.
REFERENCE	1 (bases 1 to 10643) Maeder, D.L., Weiss, R.B., Dunn, D.M., Cherry, J.L., Gonzalez, J.M., DiRuggiero, J., and Robb, F.T. Divergence of the hyperthermophilic archaea Pyrococcus furiosus and P. horikoshii inferred from complete genomic sequences Genetics 152 (4), 1299-1305 (1999)
JOURNAL	Genetics 152 (4), 1299-1305 (1999)
MEDLINE	99359404
PUBMED	10430560
REFERENCE	2 (bases 1 to 10643) Robb, F.T., Maeder, D.L., Brown, J.R., DiRuggiero, J., Stump, M.D., Yeh, E.K., Weiss, R.B., and Dunn, D.M. Genomic sequence of hyperthermophile, Pyrococcus furiosus: Implications for physiologic and enzymology Meth. Enzymol. 330, 134-157 (2001)
JOURNAL	Meth. Enzymol. 330, 134-157 (2001)
MEDLINE	21079003
PUBMED	11210495
REFERENCE	3 (bases 1 to 10643) Weiss, R.B., Dunn, D.M., Robb, F.T. and Brown, J.R. The complete sequence of the Pyrococcus furiosus genome Unpublished
JOURNAL	
MEDLINE	
PUBMED	
REFERENCE	4 (bases 1 to 10643) Weiss, R.B. Direct Submission Submitted (12-FEB-2002) Human Genetics, University of Utah, 20 South 2030 East, Salt Lake City, UT 84112, USA
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	Location/Qualifiers 1..10643 /organism="Pyrococcus furiosus DSM 3638" /mol_type="genomic DNA" /strain="DSM 3638" /db_xref="taxon:186497" complement (68..1261) /genes="PF0471" complement (68..1261) /genes="PF0471" /notes="Function Code: 16.1 Conserved Hypothetical" /codon_start=1 /translation="11 /product="hypothetical protein" /protein_id="AAL80595.1" /db_xref="GI:18892446" /translation="MIPLILGILVILTSYTDIKTSYIYEBRFPFSLSKWCKR KGVNVPKPYPIVEAGILYVFFLIGKNEILILAFSGIIGLLGFAVGSLLVYVGGW ASGDVILIAFSLILPYAPESAKYKPAVAVLPLNATILFNSILILFELIYVAFAG LIAKKLSLLKELFVKGINRVFVSLWIMFSTLTFVVISINIGLSLHPLLRWILFLL IAFLKIRVIGDILGVLSLVYVILVGFSSLYFLRLVILLYAFKILLSVSKMRKEV LTKRVREKEGELVIGERILVGVKIVDRKDFEKLNEVLKSGKEFGISKEIAGF SVGLTKEQIEELKLVSRGKLENEFLVRKAMPFALFLGLVLTYSYFFGDLWILK VSGMV" complement (1301..1438) /genes="PF0472" complement (1301..1438) /genes="PF0472" /notes="Function Code: 16.4 Hypothetical" /codon_start=1 /translation="11 /product="hypothetical protein" /protein_id="AAL80596.1" /db_xref="GI:18892447" /translation="MKRAQTAIEYLLMLAAVILVAIVLVVVDSMRTLTNYVESQK S" 1492..2172 /genes="PF0473" 1492..2172

```

/genes="PF0473"
/notes="Function Code: 13.2 Translation: Degradation of
proteins, peptides, and glycopeptides; (gcp-2)"
/codon_start=1
/transl_table=11
/product="o-sialoglycoprotein endopeptidase"
/protein_id="AAL80597.1"
/db_xref="GI:18892448"
/translacion="MKGELMKLIQGAELKYLADFSELYDLPIKIVVERSKRY
RIPEDIKERTREARILHRAKAGVYVFEVDKMIIVWSEIEGERLKELL
EKLPIDIERIEIKETGRVIGKLEHAGIVHGLDTNMMRDGKIYLIIDFGLAEFDDTL
EAQGVDLHLRLRAMESTHYSWVEGFKAVLRGYEVRGKREIEEKIKIELRGY
RERSMTS"
complement (2162..3850)
/genes="PF0474"
complement (2162..3850)
/genes="PF0474"
/notes="Function Code: 11.1 Replication: DNA replication,
restriction, modification, recombination, and repair;
(mutS)"
/codon_start=1
/transl_table=11
/product="DNA mismatch repair protein"
/protein_id="AAL80598.1"
/db_xref="GI:18892449"
/translacion="MKLRGDAREVYKRLSLRSMIKLGEARTFLKKEPFTSDREBII
KQNYLKEGLKQVRDDLELYLSIRPFRREFPHDRILLVSDSEVEAEKVLDPVPT
SPSEIEDYPLLIISTIGYIEVEVKPSHTAPELYIIPWENEDVLEALSVPFGAAD
KLISUKIEBIFPKMEILENDEIIVEKEKELNRIKIEKLPKLTISGRDLVEFMK
ALRAGELVLFHFKSALNDEIIEINKAEKESIDVLGISEVIFPDFFVEVPEQIEA
LKELEBEFKIEFYLSRSTVEKLIPLHQLKKEEIKAYELYFLVLCVSGKTSLEL
IVEEGIFTEGNLFTENPQVSYFVFKSYGNPFGVEANIVILTCAVSGKTSLEL
ISOVILAHGWPVPAKAAFTVLOEIFFKRSVYGAGAFETSLKGLNVAIKGKGK
KLILDEFESIPEGAAXIKLAEALLIAVEKGFVVIVSHLGEDLKRIBIPARVDIE
AKGLDENLILVDROPKFGVIGRSPELILVERLARKGRGEKQWNRILKXFRK"
complement (3847..4677)
/genes="PF0475"
complement (3847..4677)
/genes="PF0475"
/notes="Function Code: 13.6 Translation: Translation
factors; (eIF2B)"
/codon_start=1
/transl_table=11
/product="translation initiation factor eIF-2b, subunit
delta"
/protein_id="AAL80599.1"
/db_xref="GI:18892450"
/translacion="MLPEKVVKILEMKNEKIRGASWMAKGAEPILLSBELDSTSL
EGEIIELKREILEINSMASLYNLAMFIPITNDREVVKLRAEEFIKRAEEAKETASI
GAQIDSGDVIITHSSAVFELKTAGRGKQFVILTESAPDEGLYLAALQDES
IEVEIITDQGLGFAKDALVAGDVTVDGTVVKNAGTYLLAISCVSESEVFFVAA
ETVYKQKITSKEVLEVERPLYREGSRVNLFDITPWKPIRGIIITELGILPPRDMI"
complement (4677..5435)
/genes="PF0476"
complement (4677..5435)
/genes="PF0476"
/notes="Function Code: 16 Unknown:"
/codon_start=1
/transl_table=11
/product="competence-damage inducible protein, putative"
/db_xref="GI:18892451"
/translacion="MPAEIITVGDDELLTGNTVNSNAHIAKLTGRLVILKRIITVGD
DVSDIRSVILELNRPEDVLLIISGGIGPHDDVTWVAVALAKGELVLECKEIEIKE
FYRLRYREGIIDDPTDRLKAIPLKPAIGVAKGAPGAYIEYGTIKFVLPGM
PRMKAMLEEVLEPRIGSKFKVQKYLAEITDESKLAPLIEETIKIFDVKHSPPKGF
KFGIIGIIPAOEDDTIKKAVEFMKSGIEFREGW"
5547..6968
/genes="PF0477"
complement (5547..6968)
/genes="PF0477"
/notes="Function Code: 5.1 Central Intermediary Metabolism:

```

```

Degradation of polysaccharides"
/codon_start=1
/transl_table=11
/product="alpha-amylase"
/protein_id="AAL80601.1"
/db_xref="GI:18892452"
/translacion="MIHYETLVYGGDHVNIKKLTPLLTLLFFIVLASPVSAAKYLEL
EEGVLMAQPVWDVPGGIGWMDHRSKIPEWYEGISA.WLPPSPGMSGGYSGMYDP
YDFDGEYVQKGTETRSQSKBELVRLIQTAHAYGIVKVIADVINHRAGGLEWNPFF
YDITDTSKSVASGKITANYLDFHNEHLCCDEGTFGFPDICHKKEWQVWLKSN
ESAAALYRIGFDGFRDYVYKGAWVRDMMWGWAVGKATFVNDALLSWAYES
GAKVDFPYLYKDEAFDNNNIIPALYALQNGQTVVRDPRFKAATFVNDALLSWAYES
YPAYAFILYEGQVFIYRDFEELNKKLILNIWIHDLHLAGGSTIIVYDNDLFFV
RNGDSRRPGLITVINLSPNNVGRWVYVPKPAAGACIHEYTGNLGGWVDKRVDSGGWYL
EAPHPDPANGYCYGYSVYSGVCG"
complement (6894..8966)
/genes="PF0478"
complement (6894..8966)
/genes="PF0478"
/notes="Function Code: 5.1 Central Intermediary Metabolism:
Degradation of polysaccharides"
/codon_start=1
/transl_table=11
/product="alpha-amylase"
/protein_id="AAL80602.1"
/db_xref="GI:18892453"
/translacion="MVDRFVDRNPTNNEPFYDPEKKNRYLYMGDDIEGIIERLDYIES
LGVSWLSPLNDNINRMAGSAPYHYWPRDFKRIDEHFGTWEDFRRLVEEAKKRG
CIIVDVVPHNSPATDGEFGALYDNGTLVTVNYEDKNATRNPTATSLENIYHNGNI
NDNFQQLKXANLFGLDNFQNDVDVNLKEGALFVKNAGCGFRIDAKHIELGWL
ETFLYDIQSDPEFLIYGFYFANTPDKTFLDYEFYRYSNVSSLLNIPRESIARFPA
YGGFQQLAKMEEYYSLVFVYKQNLDFDSDHLDFVLRNPNKORIRHAGLGLWLP
GIVYVYDESILYKSGKDPYRPMVMDNSTKAEIIRKLSLRKVDALAYSDP
RTVVDYNTWIFERKFGSHKILVAKGPDKNITISLNTDSTYDIDIEGAILKVEG
QGIKIPRYGFYVHVEEOKTPLIGITPYIAQPGQKILLAGLNGNIKVYIGGR
ARIIEKENSILVEPEIKTMNAWIPWVWNGTRSNVKKLYRYSNDNIPALIVLQGN
YTGLVWKNIPLESSEPLKSPTHGYFAVPLPBNKFTVQLYKGLWPELOPTNV
TLYGIGNKTVILNEAAPETPVCGFIVAFVALLPLLRKKKKSPNTTITPYGVPIVAV
SWIVRMCL"
complement (9137..10534)
/genes="PF0479"
complement (9137..10534)
/genes="PF0479"
/notes="Function Code: 16.4 Hypothetical"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL80603.1"
/db_xref="GI:18892454"
/translacion="MPKKLLVAVLFLFLLLSLLGRRYLGAFFKPKLOFILVDENGT
PLKISKNVQVQIDALVPTDEVYKTYIGTLKKPTILKFWDSGNLIKTSVADQKIQ
SVISKWVEQSEISGFSVISVWVLDYENKLRGTIIYNYEIKDMRKGFKKIVKYNL
HKLPSSELPNRTASIQSGSNKQFYKVKVDSRWSRPNNDYIKIPILIVDNKYSSEGLIN

```

Query Match 60.5%; Score 792.8; DB 1; Length 10643;

Best Local Similarity 75.4%; Pred. No. 1.4e-168;

Matches 986; Conservative 0; Mismatches 322; Indels 0; Gaps 0;

Qy	4	GCCAAAGTACTCCGAGCTGGAAAAGCGCGGGTCTAATAGAGCGGCTTCTACTGGAGCTG	63
Db	5661	GCAAAATACTTGGAGCTTGAAGAGGAGGAGTATATATGCAAGCATTTCTATTGGGATGT	5720
Qy	64	CCTTCAGGAGGAATATGTTGGGACACAAATACGGCAGAGATACCGAGTGGTACGATGCC	123
Db	5721	CCAGGGGAGGAATTTGGTGGGATCATATAGATCGAGATTCTGATGATGAGCT	5780
Qy	124	GGAAATCTCCGCAATATCGATTCCCCCGGCGAGCAAGGGCATTGGCGCGCCCTATTGATG	183
Db	5781	GGAAATCTCTCAATATGGTACCTCCACCAAGCAAGGGGATGATGGAGGATATTCAATG	5840
Qy	184	GGCTACGACCCCTACGACTTCTTTCCTCGGTGAGTACGACCAAGGAGGAAACGTTAGAG	243
Db	5841	GGCTACGATCCCTATGATTTACTTTTGATCTCGGAGGACTTACCAAGAGGAACTGTAGAG	5900



Search completed: June 29, 2004, 10:43:39  
Job time : 5333 secs

```
Qy 244 ACGCGCTTGGCTCCAGCAGGAGCTCGTGAAACATGATAAACAACCGGCCACGCGCTATGGC 303
Db |||||
Db 5901 ACGCGCTTGGATCAAAAGAGAGACTAGTGAGATTGATACAAACTGCCATGCCCTATGGA 5960
Qy 304 ATGAAGGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGGAGTGGAAAC 363
Db |||||
Db 5961 ATAAGGTAATCGCGGATGTAGTTATTAACACACAGGCGTGTGTGACCTAGATATGGAAC 6020
Qy 364 CCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAAGGTGCGGTGCGGTAAATACAG 423
Db |||||
Db 6021 CCCTTCGTGGAGATTACATGACACAGACTTTTCTAAAGTTGCTCAGGGAATATACA 6080
Qy 424 GCCAATACCTCGACTTCCACCCGAAAGAGCTCCATCGCGCGGATTCGGGAACATTTGGA 483
Db |||||
Db 6081 GCTAACTATCTGGACTTCCATCCAAAGAGCTTCATTTGTTGTGACGAAGGAACCTTTGGA 6140
Qy 484 GGCTATCCCGACATATGCCAGCAAGAGCTGGACCCAGTACTGGCTCTGGGCCAGCCAG 543
Db |||||
Db 6141 GGATTTCCAGATATATGTCATCACAAAGAGTGGGATCAGTACTGGCTATGGAAGAGCAAT 6200
Qy 544 GAGAGCTACGGGGCATATCTCAGAGAGCATCGGCATCGATCGCTGGCGCTTCGACTACGTC 603
Db |||||
Db 6201 GAGAGTTATGCTGCTTATTTAAGAGCATAGGATTTGATGGTTGGAGATTGACTATGTT 6260
Qy 604 AAGGCTATGCTCCCTGGTCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGCGGTT 663
Db |||||
Db 6261 AAGGCTATGAGCTTGGGTTGTGAGAGCTGGCTTAATTTGGTGGGAGGTTGGGCAATT 6320
Qy 664 GGAGAGTACTGGGACACCAACGTGCAAGCTGTTCTCAACTGGGCACTCTGAGCGGTGCC 723
Db |||||
Db 6321 GGAGAGTACTGGGACACAAATGTAGATGCACTACTAAGCTGGGCATATGAGAGTGTGCA 6380
Qy 724 AAGGCTTTGACTTGGCTCTACTACAGATGGATGAGGCTTTGACACAAACAACTT 783
Db |||||
Db 6381 AAGGCTTTGACTTGGCTCTACTATAAATGGATGAGCAATTTGACAAATACAACTT 6440
Qy 784 CCAGCGCTCGTCTCTGCCCTTCAGAACCGCCAGACTGTTGTCTCCCGCAGCCCGTTCAAG 843
Db |||||
Db 6441 CCAGCATTTAGTCTATGCCCTACAAACGACAACTGTAGTTTCGAGAGATCCATTTAAG 6500
Qy 844 GCCGTAACTTTGTAGCAAAACACAGACCGATATATCTGGAACAGTATCCAGGCTAC 903
Db |||||
Db 6501 GCAGTAACCTTTCGTTGCCAATCATGACACAGATATAATATGGAACAAAGTATCCAGCATAT 6560
Qy 904 GGGTTTCATCTCACCTAGAGGCGCAGCGCAATATTTCTACCGGACTACGAGGAGTGG 963
Db |||||
Db 6561 GGGTTTCATATTGACATATAGGGACAGCCAGTATATTTCTACAGGACTTTGAGGAATGG 6620
Qy 964 CTCACAAAGGATPAGCTCAAGAACTCATCTGGATACATGAGAACCTCGCCGGAGGAAGC 1023
Db |||||
Db 6621 CTGAACAAGGATAGCTAATTAACCTCATTTGGATCCATGATCATTTGSCAGAGGAAGC 6680
Qy 1024 ACCGACATAGTCTACTACGATACGATGAATCTATCTTCGTCAGGAACGGCTACGGGAC 1083
Db |||||
Db 6681 ACAACAAATGTTCTACTACGACACCATGAGTCAATTTGTGAGAAATGGAGATTTCTAGA 6740
Qy 1084 AAGCGGGGCTTTAACTTACCTTACATCACTAGCTCGAGCAAGCGCGGAAGGTGGTTTAT 1143
Db |||||
Db 6741 AGGCTTGGGCTTTAACTTACATTTACATTTAGCTTGAAGCTTGAAGTGGGTATAC 6800
Qy 1144 GTGCGGAAGTTCCGGGCGCGGTGATCCACGAGTATAGTGAACCTCGAGGCTGGGTA 1203
Db |||||
Db 6801 GTTCCAAAGTTTGCAGGGGCTTGTATTATGAAATACATGGAACCTAGGAGGATGGGTA 6860
Qy 1204 GACAGTACGCTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCTGCC 1263
Db |||||
Db 6861 GATAAAGAGTAGATAGTAGCGGATGGGTATACCTAGAGGACCACTTACGATCCAGCT 6920
Qy 1264 AACGGGCACTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGGCTGA 1311
Db |||||
Db 6921 AACGGCTACTATGGGTACTCCGCTATGGAGTTATTGTGGTGTGGGTGA 6968
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 09:04:55 ; Search time 573 Seconds  
(without alignments)  
9719.699 Million cell updates/sec

Title: US-10-081-739A-1  
Perfect score: 1311  
Sequence: 1 atgscgaagtactccgagct.....gctactgcgggtggctga 1311

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3173863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_29Jan04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002s:\*
  - 7: Geneseqn2003as:\*
  - 8: Geneseqn2003bs:\*
  - 9: Geneseqn2003cs:\*
  - 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1311	7	ABX08425 DNA encod
2	1311	100.0	1311	7	ABX08420 DNA encod
3	1239	94.5	1311	7	ABX08414 DNA encod
4	1231	93.9	1311	7	ABX08442 DNA encod
5	1192.6	91.0	1311	7	ABX08448 DNA encod
6	1186.2	90.5	1311	7	ABX08426 DNA encod
7	1183	90.2	1311	7	ABX08428 DNA encod
8	1171.8	89.4	1311	7	ABX08429 DNA encod
9	1160.6	88.5	1311	7	ABX08424 DNA encod
10	1155.8	88.2	1311	7	ABX08416 DNA encod
11	1147.8	87.6	1311	7	ABX08440 DNA encod
12	1147.8	87.6	1311	7	ABX08437 DNA encod
13	1146.2	87.4	1311	7	ABX08422 DNA encod
14	1146.2	87.4	1311	7	ABX08432 DNA encod
15	1146.2	87.4	1311	7	ABX08412 DNA encod
16	1143	87.2	1311	7	ABX08439 DNA encod
17	1141.4	87.1	1311	7	ABX08435 DNA encod
18	1139.8	86.9	1311	7	ABX08441 DNA encod
19	1138.2	86.8	1311	7	ABX08419 DNA encod
20	1130.2	86.2	1311	7	ABX08417 DNA encod
21	1120.6	85.5	1311	7	ABX08415 DNA encod
22	1119	85.4	1311	7	ABX08430 DNA encod
23	1117.4	85.2	1311	7	ABX08411 DNA encod
					ABX08427 DNA encod
					ABX08421 DNA encod

ALIGNMENTS

RESULT 1

ID ABS55928 standard; DNA; 1311 BP.

XX

AC ABS55928;

DT 20-JAN-2003 (first entry)

XX

DE DNA encoding alpha amylase polypeptide #1.

XX

KW Alpha amylase; gene; ds; starch liquefaction; starch linkage hydrolysis;  
KW starch breakdown catalysis; high-maltose syrup; high-glucose syrup; rice;  
KW germinated rice; corn; barley; wheat; legume; sweet potato; anti-staling;  
KW textile desizing; lignocellulosic fibre treatment; enzymatic de-inking;  
KW recycled paper pulp; alpha-glucosidic linkage; dough; crumb firming;  
KW crumb elasticity; palatability; improved flavour.

Unidentified.

OS

XX

XX

Key

Location/Qualifiers

1..1311

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

WO200268597-A2.

06-SEP-2002.

21-FEB-2002; 2002WO-US005538.

21-FEB-2001; 2001US-0270495P.

21-FEB-2001; 2001US-0270496P.

14-MAY-2001; 2001US-0291122P.

(DIVE-) DIVERSA CORP.

Callen W, Richardson T, Frey G, Miller C, Kazaoka M, Mathur EJ;

Short JW;

WPI; 2003-018657/01.

P-PSDB; ABG70531.

Novel alpha amylase polypeptide, useful e.g. for starch liquefaction,

textile desizing, treating lignocellulosic fibers and for preparing dough

or baked product, has increased activity and stability.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PS Claim 1; Fig 7; 147pp; English.

XX The invention relates to a purified alpha amylase polypeptide and the  
 CC polynucleotide encoding it. The polypeptide is useful for liquefying a  
 CC starch containing composition, hydrolysing a starch linkage, catalysing  
 CC the breakdown of a starch and producing high-maltose or high-glucose  
 CC syrup or mixed syrup, where the starch is from rice, germinated rice,  
 CC corn, barley, wheat, legumes or sweet potato. The polypeptide is also  
 CC useful for washing an object, textile desizing, treating lignocellulosic  
 CC fibres for improving fibre properties, enzymatic de-inking of recycled  
 CC paper pulp and modifying small molecules by biocatalytic reaction. The  
 CC polypeptide is further useful for increasing the flow of production  
 CC fluids from a subterranean formation by removing a viscous, starch-  
 CC containing, damaging fluid formed during production operations and found  
 CC within the subterranean formation which surrounds a completed well bore,  
 CC by pumping an enzyme treatment comprising the polypeptide and an aqueous  
 CC fluid to a desired location within the well-bore which is effective to  
 CC attack the alpha-glucosidic linkages in the starch-containing fluid. The  
 CC alpha amylases have increased activity and stability at increased pH and  
 CC temperature and can be used in preparing dough to provide an improved  
 CC anti-staling effect as measured by less crumb firming, retained crumb  
 CC elasticity, improved slice-ability, improved palatability or improved  
 CC flavour. This sequence represents DNA encoding an alpha amylase of the  
 CC invention

SQ Sequence 1311 BP; 318 A; 364 C; 376 G; 253 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 7; Length 1311;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCAACTACCGAGTGGAAAGCGGGGTCATATCGAGCGCTTCTACTGGGAC 60  
 DB 1 ATGGCCAACTACCGAGTGGAAAGCGGGGTCATATCGAGCGCTTCTACTGGGAC 60

QY 61 GTCCCTTCAGGAGGAATATGGTGGGACACAAATACGGCAGAAATACCGAGTGGTACGAT 120  
 DB 61 GTCCCTTCAGGAGGAATATGGTGGGACACAAATACGGCAGAAATACCGAGTGGTACGAT 120

QY 121 GCGGGAATCTCCGCAATATGGATTCGCCGGGAGCAAGGGATCGGGCGGCTATTGCG 180  
 DB 121 GCGGGAATCTCCGCAATATGGATTCGCCGGGAGCAAGGGATCGGGCGGCTATTGCG 180

QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACACAGAGGCAACGGTA 240  
 DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACACAGAGGCAACGGTA 240

QY 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAAACACCGCCCTAT 300  
 DB 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAAACACCGCCCTAT 300

QY 301 GGCATGAAGTAAATAGCCGATATAGTCATCACCCGCGCGCGGTGACCTGGAGTGG 360  
 DB 301 GGCATGAAGTAAATAGCCGATATAGTCATCACCCGCGCGCGGTGACCTGGAGTGG 360

QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTGCGCTCGGGTAAATAC 420  
 DB 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTGCGCTCGGGTAAATAC 420

QY 421 ACGGCCAACTACCTCGACTTCACCCGAGAGCTCCATCGCGGCGATTCGGACATTT 480  
 DB 421 ACGGCCAACTACCTCGACTTCACCCGAGAGCTCCATCGCGGCGATTCGGACATTT 480

QY 481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGGACCACTACTGGCTCTGGGCCAGC 540  
 DB 481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGGACCACTACTGGCTCTGGGCCAGC 540

QY 541 CAGGAGAGCTACCGGGCATATCTCAGAGGATCGGATCGATCGCTGGGCTTCGACTAC 600  
 DB 541 CAGGAGAGCTACCGGGCATATCTCAGAGGATCGGATCGATCGCTGGGCTTCGACTAC 600

QY 601 GTCAAGGGCTATCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
 DB 601 GTCAAGGGCTATCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660

DB 601 GTCAAGGGCTATCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
 QY 661 GTTGGAGAGTACTGGGACACCAACAGCTCGACCGCTGTCTCAACTGGGCATATCTCGAGCGGT 720  
 DB 661 GTTGGAGAGTACTGGGACACCAACAGCTCGACCGCTGTCTCAACTGGGCATATCTCGAGCGGT 720  
 QY 721 GCCAAGGCTTTTACTTGGCCCTCTACTACAAGATGGATGAGGCTTTTACCAACAAAAAC 780  
 DB 721 GCCAAGGCTTTTACTTGGCCCTCTACTACAAGATGGATGAGGCTTTTACCAACAAAAAC 780  
 QY 781 ATTCCAGCGCTGCTCTGCTGCTTCAAGAGCGGACAGCTGTGTCTCCCGCCACCCGCTTC 840  
 DB 781 ATTCCAGCGCTGCTCTGCTGCTTCAAGAGCGGACAGCTGTGTCTCCCGCCACCCGCTTC 840  
 QY 841 AAGGCGGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAAACAGTATCCAGCC 900  
 DB 841 AAGGCGGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAAACAGTATCCAGCC 900  
 QY 901 TAGCGGTTCATCTCCTACCTACGAGGCGCCAGCCGACATATTTACCGGACTACGAGGAG 960  
 DB 901 TAGCGGTTCATCTCCTACCTACGAGGCGCCAGCCGACATATTTACCGGACTACGAGGAG 960  
 QY 961 TGGCTCAACAAAGGATAAGCTCAAGAAACCTCATCTGGATACATGAGAACTCTGCGCGGAGGA 1020  
 DB 961 TGGCTCAACAAAGGATAAGCTCAAGAAACCTCATCTGGATACATGAGAACTCTGCGCGGAGGA 1020  
 QY 1021 AGACCGGATAGTCTACTACGATACGATAGTAACTCATCTTCTGTCAGGACGCTACGGG 1080  
 DB 1021 AGACCGGATAGTCTACTACGATACGATAGTAACTCATCTTCTGTCAGGACGCTACGGG 1080  
 QY 1081 GACAAGCGGGGCTTATAACTACATCAACTAGGCTCGAGCAAGCCGGAAGCTGGGT 1140  
 DB 1081 GACAAGCGGGGCTTATAACTACATCAACTAGGCTCGAGCAAGCCGGAAGCTGGGT 1140  
 QY 1141 TATGTGCGAAGTTCGCGGCGGTGATCCAGTATCTGTTAACTTCTGAGGCTGG 1200  
 DB 1141 TATGTGCGAAGTTCGCGGCGGTGATCCAGTATCTGTTAACTTCTGAGGCTGG 1200  
 QY 1201 GTAGCAAGTACGCTTACTCAAGCGGTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 DB 1201 GTAGCAAGTACGCTTACTCAAGCGGTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 QY 1261 GCCAAGCGGCGAGTATGCTACTCCGTGTGAGTACTTGGGGGTGGGCTGA 1311  
 DB 1261 GCCAAGCGGCGAGTATGCTACTCCGTGTGAGTACTTGGGGGTGGGCTGA 1311

RESULT 2  
 ABZ23371  
 ID ABZ23371 standard; DNA; 1311 BP.  
 XX  
 AC ABZ23371;  
 XX  
 DT 07-APR-2003 (first entry)  
 XX  
 DE Nucleotide sequence of an alpha-amylase.  
 XX  
 KW Alpha-amylase; enzyme; detergent; baking; beverage; fuel; ethanol; gene;  
 KW ss.  
 XX  
 OS Bacillus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1311  
 FT /\*tag= a  
 FT /product= "alpha-amylase"  
 XX  
 XX WO200292802-A1.  
 XX  
 XX 21-NOV-2002.  
 XX  
 XX 14-MAY-2002; 2002WO-US015422.  
 XX

PR 14-MAY-2001; 2001US-0291122P.  
XX (DIVE-) DIVERSA CORP.  
FA Gerendash J;  
XX WPI; 2003-156728/15.  
DR P-PSDB; ABP70154.  
XX  
XX Purifying enzyme, such as alpha-amylase and cellulase, involves  
PT flocculating fermentation broth containing bacterial cells containing the  
PT enzyme, centrifuging and extracting enzymes by filtration.  
PS Disclosure; Fig 7; 108pp; English.  
XX  
CC The present sequence encodes an alpha-amylase enzyme, which may be  
CC purified using the method of the invention. The specification describes a  
CC method for purifying an enzyme, such as alpha-amylase and cellulase, by  
CC flocculating fermentation broth containing bacterial cells containing the  
CC enzyme, centrifuging and extracting enzymes by filtration. Alternatively,  
CC the above method involves subjecting a fermentation broth to a heat-  
CC killing procedure, where a resultant heat-killed broth is formed, washing  
CC the heat-killed broth with a buffered medium, releasing the enzyme  
CC contained in the cells of the fermentation broth and extracting the  
CC enzyme by microfiltration. The method is useful for purifying an enzyme  
CC derived from a mixed population of organisms or an isolate from the  
CC population. The alpha-amylases are useful in corn-wet milling processes,  
CC detergents, baking processes, beverages and in oil fields (fuel ethanol)  
XX  
SQ Sequence 1311 BP; 318 A; 364 C; 376 G; 253 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 7; Length 1311;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCCAGTACTCGAGCTGGAAAGGGCGGGTCAATATGCAAGCGTTCTACTGGAC 60  
Db 1 ATGGCCAGTACTCGAGCTGGAAAGGGCGGGTCAATATGCAAGCGTTCTACTGGAC 60  
Qy 61 GTGGCTTCAGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120  
Db 61 GTGGCTTCAGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120  
Qy 121 GCCGGAATCTCCGAATATGATTTCCCGCGAGCAAGGCGATGCGCGCGCTATTTCG 180  
Db 121 GCCGGAATCTCCGAATATGATTTCCCGCGAGCAAGGCGATGCGCGCGCTATTTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
Qy 241 GAGACGGCTTTGGCTCCAGCAGAGCTCGTGAACATGATAAACACCGCCACGCCCTAT 300  
Db 241 GAGACGGCTTTGGCTCCAGCAGAGCTCGTGAACATGATAAACACCGCCACGCCCTAT 300  
Qy 301 GGCATGAAGTAAATAGCGATATAGTCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGTAAATAGCGATATAGTCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTTCGGTGGGTAATAC 420  
Db 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTTCGGTGGGTAATAC 420  
Qy 421 ACGGCCAACTACCTCGACTTCCACCGGAAACGAGTCCATCGTGGCGGATTCGGAAACATTT 480  
Db 421 ACGGCCAACTACCTCGACTTCCACCGGAAACGAGTCCATCGTGGCGGATTCGGAAACATTT 480  
Qy 481 GGAGGCTATCCGACATATGCGACAGAGCTGGGACCATGATCTGGCTCTGGGCCAGC 540  
Db 481 GGAGGCTATCCGACATATGCGACAGAGCTGGGACCATGATCTGGCTCTGGGCCAGC 540  
Qy 541 CAGGAGCTACCGCGCATATCTCAGAGCATCGGCATCGCTCGCGCTTCGACTAC 600  
Db 541 CAGGAGCTACCGCGCATATCTCAGAGCATCGGCATCGCTCGCGCTTCGACTAC 600

Db 541 CAGGAGCTACCGCGCATATCTCAGAGCATCGGCATCGCTCGCGCTTCGACTAC 600  
Qy 601 GTCAAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAACCTGGTGGGAGCGTGGCG 660  
Db 601 GTCAAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTGAACCTGGTGGGAGCGTGGCG 660  
Qy 661 GTTGGAGAGTACTGGGACACCAACCTGCAAGCGTGTCTCAACTGGGCGATCTCGAGCGGT 720  
Db 661 GTTGGAGAGTACTGGGACACCAACCTGCAAGCGTGTCTCAACTGGGCGATCTCGAGCGGT 720  
Qy 721 GCCAAGGTCTTTGACTTGGCCCTCTACTACAAGATGGATGAGGCTTTGACACCAAAAAA 780  
Db 721 GCCAAGGTCTTTGACTTGGCCCTCTACTACAAGATGGATGAGGCTTTGACACCAAAAAA 780  
Qy 781 ATTCCAGCGCTCGTCTCTGCTCAGAACGGCGAGCTGTGTCTCCGCGGACCGGCTC 840  
Db 781 ATTCCAGCGCTCGTCTCTGCTCAGAACGGCGAGCTGTGTCTCCGCGGACCGGCTC 840  
Qy 841 AAGGGCGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
Db 841 AAGGGCGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
Qy 901 TAGCGTTCATCTCTACCTACGAGGCGCGAGCGGACAAATATTCTACCGGACTACGAGGAG 960  
Db 901 TAGCGTTCATCTCTACCTACGAGGCGCGAGCGGACAAATATTCTACCGGACTACGAGGAG 960  
Qy 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGATACATGAGAACCTCGCCGAGGA 1020  
Db 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGATACATGAGAACCTCGCCGAGGA 1020  
Qy 1021 AGCACCGCATAGTCTTACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG 1080  
Db 1021 AGCACCGCATAGTCTTACGATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG 1080  
Qy 1081 GACAGCGGGGCTTATACCTACATCACTAGGCTCGAGCAAGCGGAGGAGTGGGT 1140  
Db 1081 GACAGCGGGGCTTATACCTACATCACTAGGCTCGAGCAAGCGGAGGAGTGGGT 1140  
Qy 1141 TATGTGCCAAGTTCGCGGCGCTGATCACTACAGAGTACTCTGTAACCTCGGAGGCTGG 1200  
Db 1141 TATGTGCCAAGTTCGCGGCGCTGATCACTACAGAGTACTCTGTAACCTCGGAGGCTGG 1200  
Qy 1201 GTAGACAGTACGTCTAATCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Db 1201 GTAGACAGTACGTCTAATCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Qy 1261 GCCAACGGGCGATGAGTCTGCTTCTGAGCTACTCGCGGCTGGGCTGA 1311  
Db 1261 GCCAACGGGCGATGAGTCTGCTTCTGAGCTACTCGCGGCTGGGCTGA 1311

RESULT 3  
ABX08413  
ID ABX08413 standard; DNA; 1311 BP.  
XX  
AC ABX08413;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE DNA encoding alpha amylase polypeptide #3.  
XX  
KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
OS Unidentified.  
XX  
PN WO200268589-A2.  
XX  
PD 06-SEP-2002.

XX PF 21-FEB-2002; 2002WO-US005068.  
 XX PR 21-FEB-2001; 2001US-0270495P.  
 XX PR 21-FEB-2001; 2001US-0270496P.  
 XX PR 14-MAY-2001; 2001US-0291122P.  
 XX PA (DIVE-) DIVERSA CORP.  
 XX PI Callen W, Richardson T, Frey G;  
 XX DR WPI; 2003-018656/01.  
 XX DR P-FSDB; AB003042.  
 XX PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX PS Claim 1; Fig 16; 30pp; English.  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel industry), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention  
 XX SQ Sequence 1311 BP; 312 A; 367 C; 375 G; 257 T; 0 U; 0 Other;

Query Match 94.5%; Score 1239; DB 7; Length 1311;  
 Best Local Similarity 96.6%; Pred. No. 0;  
 Matches 1266; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATGGCCAACTCTCCGAGCTGGAAAGGGCGGGTCATAATGCAGGCGCTTCTACTGGGAC 60  
 DB 1 ATGGCCAACTCTCCGAGCTGGAAAGGGCGGGTCATAATGCAGGCGCTTCTACTGGGAC 60  
 QY 61 GTGCCCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGTCAGAT 120  
 DB 61 GTGCCAGGTGGAGGAATCTGTGGGACACCATCAGGAGCAGATACCGGAGTGTCAGAG 120  
 QY 121 GCGGGAATCTCGCAATATGGATTCGCCCGGAGCAGGAGGATCGGCGGCGCTATTTCG 180  
 DB 121 GCGGGAATATCGGCCATTTGGATTCGCCCGGAGCAGGAGGATCGGCGGCGCTATTTCG 180  
 QY 181 ATGGGGTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACACGAGGAGACGGTA 240  
 DB 181 ATGGGGTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACACGAGGAGACGGTA 240  
 QY 241 GAGACCGCTTTGGCTTCCAGCAGGAGCTCGTGGAACATGATATAACACCGCCACGCTAT 300  
 DB 241 GAGACCGCTTTGGCTTCCAGCAGGAGCTCGTGGAACATGATATAACACCGCCACGCTAT 300  
 QY 301 GGCATGAAGGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGAGTGAGTG 360  
 DB 301 GGCATGAAGGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGAGTGAGTG 360  
 QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTCGCGGTAAATAC 420  
 DB 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTCGCGGTAAATAC 420

QY 421 ACGGCCAACTACCTCGACTTCCACCAGAACGAGCTCCATCGGGGCGATTCCGGAAACATTT 480  
 DB 421 ACGGCCAACTACCTCGACTTCCACCAGAACGAGCTCCATCGGGGCGATTCCGGAAACATTT 480  
 QY 481 GGAGGCTATCCCGACATATGCCACGACAAAGAGCTGGGACAGTACTTGGCTCTGGGGCAGC 540  
 DB 481 GGAGGCTATCCCGACATATGCCACGACAAAGAGCTGGGACAGTACTTGGCTCTGGGGCAGC 540  
 QY 541 CAGGAGAGCTACCGGCGCATATCTCAGGAGCATCGGATCGATCGATCGATCGATCGATCG 600  
 DB 541 CAGGAGAGCTACCGGCGCATATCTCAGGAGCATCGGATCGATCGATCGATCGATCGATCG 600  
 QY 601 GTCAAGGGCTATGCTCCCTGGGTCTGCTCAAGGCTGGCTGAATGGTGGGGAGGCTGGGG 660  
 DB 601 GTCAAGGGCTATGCTCCCTGGGTCTGCTCAAGGCTGGCTGAATGGTGGGGAGGCTGGGG 660  
 QY 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGCATATCGAGCGGT 720  
 DB 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGCATATCGAGCGGT 720  
 QY 721 GCCAAGGTCTTTGACTTCGGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAAC 780  
 DB 721 GCCAAGGTCTTTGACTTCGGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAAC 780  
 QY 781 ATTCCAGCGCTGCTCTGCTGCTTCAAGAGGCGACACTGTTGCTCTCCCGGACCCGCTTC 840  
 DB 781 ATTCCAGCGCTGCTCTGCTGCTTCAAGAGGCGACACTGTTGCTCTCCCGGACCCGCTTC 840  
 QY 841 AAGGCGGTAACTTTGAGCAACACGACCGCATATATCTGGAACAAAGTACCTTGCT 900  
 DB 841 AAGGCGGTAACTTTGAGCAACACGACCGCATATATCTGGAACAAAGTACCTTGCT 900  
 QY 901 TAGCGCTTCATCTCACTACGAGGCGCGGACCAATATCTACCGGACTACGAGGAG 960  
 DB 901 TAGCTTTTCATCTCACTACGAGGCGCGGACCGCTCATATCTTACCGGACCAAGAGGAG 960  
 QY 961 TGGCTCAACAAAGGATAAGCTCAAGAACCTCATCTGGGATACATGAGAACCTCCCGGAGGA 1020  
 DB 961 TGGCTCAACAAAGGACAGGTTGAACAACTCATATGATACACGACCACCTCGCAGGTGA 1020  
 QY 1021 AGACCGCATAGTCTACTACGATACGATGACTCATCTTCTGTCAGGAACGCTACGGG 1080  
 DB 1021 AGACCGCATAGTCTACTACGATACGATGACTCATCTTCTGTCAGGAACGCTACGGG 1080  
 QY 1081 GACAAAGCGGGGCTTAACTCACTCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGGTT 1140  
 DB 1081 GACAAAGCGGGGCTTAACTCACTCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGGTT 1140  
 QY 1141 TATGTGCGGAAGTTTCGGGGCGGTGCATCCAGAGTATCTGGTAACTCTCGAGGCTGG 1200  
 DB 1141 TATGTGCGGAAGTTTCGGGGCGGTGCATCCACAGTATCTGGTAACTCTCGAGGCTGG 1200  
 QY 1201 GTAGACAGAGTACCTCTACTCAAGCGCTGGGTCTATCTGGAAGCTCCAGCTTACGACCT 1260  
 DB 1201 GTAGACAGAGTACCTCTACTCAAGCGCTGGGTCTATCTGGAAGCTCCAGCTTACGACCT 1260  
 QY 1261 GCCAAGCGGCGAGTATGGCTACTCCGTGTGAGTACTTGGGGGTGGGCTGA 1311  
 DB 1261 GCCAAGCGGCGAGTATGGCTACTCCGTGTGAGTACTTGGGGGTGGGCTGA 1311

## RESULT 4

ABX08445

ID ABX08445 standard; DNA; 1311 BP.

XX AC ABX08445;

XX DT 21-JAN-2003 (first entry)

XX DE DNA encoding alpha amylase polypeptide #35.

XX KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;



KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
XX starch modification.

OS Unidentified.  
XX WO200268589-A2.  
XX 06-SEP-2002.  
XX 21-FEB-2002; 2002WO-US005068.  
XX 21-FEB-2001; 2001US-0270495P.  
XX 21-FEB-2001; 2001US-0270496P.  
XX 14-MAY-2001; 2001US-0291122P.  
XX (DIVE-) DIVERSA CORP.  
XX Callen W, Richardson T, Frey G;  
XX P-PSDB; ABU03074.  
XX WPI; 2003-018656/01.  
XX P-PSDB; ABU03074.

XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
XX liquefying starch, for textile desizing, for treating lignocellulosic  
XX fibers, and for producing high-maltose or high-glucose syrup.  
XX Claim 1; Fig 16; 30pp; English.  
XX The invention relates to a purified polypeptide with alpha-amylase  
XX activity and the polynucleotide encoding it. The polypeptide is useful  
XX for hydrolysing starch linkages, for catalysing the breakdown of a  
XX starch, for modifying small molecules, for liquefying starch, for washing  
XX an object, for textile desizing, for treating lignocellulosic fibers, for  
XX improving fibre properties, for enzymatic de-inking of recycled paper,  
XX pulp, for producing a high-maltose or high-glucose syrup or a mixed  
XX syrup, and for increasing the flow of production fluids from a  
XX subterranean formation by removing a viscous, starch-containing, damaging  
XX fluid formed during production operations and found within the  
XX subterranean formation which surrounds a completed well bore. The  
XX polypeptide is also useful for preparing a dough or a baked product  
XX prepared from the dough and in corn-wet milling processes, detergents,  
XX baking processes, beverages, oil fields (fuel ethanol), brewing processes  
XX and starch modification in the paper and pulp industry, for removing  
XX starch containing stains from a material and for reducing staling of  
XX bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
XX encoding alpha amylase polypeptides of the invention  
XX Sequence 1311 BP; 322 A; 362 C; 376 G; 251 T; 0 U; 0 Other;

Query Match 93.9%; Score 1231; DB 7; Length 1311;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 1261; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 ATGGCCAGTACTCCGAGCTGGAAAGGCGGGGTTCATATGACGAGCGTTCCTACTGGGAC 60  
DB 1 ATGGCCAGTACTCCGAGCTGGAAAGGCGGGGTTCATATGACGAGCGTTCCTACTGGGAC 60  
QY 61 GTGGCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGTACGAT 120  
DB 61 GTGGCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGTACGAT 120  
QY 121 CCGGAAATCTCGCAATATGAATTCCTCCCGGAGAGCGGATGGCGGGCCCTATTTCG 180  
DB 121 CCGGAAATCTCGCAATATGAATTCCTCCCGGAGAGCGGATGGCGGGCCCTATTTCG 180  
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
QY 241 GAGAGCGGCTTGGCTCAACAGAGAGCTCTGTAAACATGATATAACACCGCCACCGCTAT 300

DB 241 GAGAGCGGCTTGGCTCAACAGAGAGCTCTGTAAACATGATATAACACCGCCACCGCTAC 300  
QY 301 GGCATGAAGGTAAATAGCCGATATAGTCATCAACACCGCGCGGGGTGAGTCACTGAGTGG 360  
DB 301 GGCATGAAGGTAAATAGCCGATATAGTCATCAACACCGCGCGGGGTGAGTCACTGAGTGG 360  
QY 361 AACCCCTTCGTGAACGACTATACCTGGAGCCGACTTCTCAAAGGTGCGGTGCGGTAAATAC 420  
DB 361 AACCCCTTCGTGGGAGACTACCTGGAGCGGACTTCTCAAAGGTAGCTCTCGGGCAATAT 420  
QY 421 AGGGCCAACTAGCTCGACTTCACCCGAAACGAGCTCCATTCGCGGCGATTCGCGAATTT 480  
DB 421 ACTGCCAACTAGCTCGACTTCACCCGAAACGAGCTCCATTCGCGGCGATTCGCGAATTT 480  
QY 481 GGAGGCTATCCGACATATGCCAGACAGAGCTGGGACCACTACTGCTGGTCTGGGCCAGC 540  
DB 481 GGAGGCTATCCGACATATGCCAGACAGAGCTGGGACCACTACTGCTGGTCTGGGCCAGC 540  
QY 541 CAGGAGAGCTAGCGGCATATCTCAGGAGCATCGGCATCGATGCTGGCGGTTCGACTAC 600  
DB 541 CAGGAGAGCTAGCGGCATATCTCAGGAGCATCGGCATCGATGCTGGCGGTTCGACTAC 600  
QY 601 GTCAAGGGCTATGCTCCCTGGGTCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGG 660  
DB 601 GTCAAGGGCTATGCTCCCTGGGTCTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGG 660  
QY 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGCATACTCGAGCGGT 720  
DB 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTCTCAACTGGGCATACTCGAGCGGT 720  
QY 721 GCCAAGGTCTTTGACTTTCGCCCTCTACTCAAGATGGATGAGGCTTTTGACAAACAAAC 780  
DB 721 GCCAAGGTCTTTGACTTTCGCCCTCTACTCAAGATGGATGAGGCTTTTGACAAACAAAC 780  
QY 781 ATTCCAGCGCTCGTCTCTGCCCTTCAGACGCGCAGACTGTTGTCTCCCGCGACCGGTTTC 840  
DB 781 ATTCCAGCGCTCGTCTCTGCCCTTCAGACGCGCAGACTGTTGTCTCCCGCGACCGGTTTC 840  
QY 841 AAGCGCGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
DB 841 AAGCGCGTAACTTTGTAGCAAAACCAACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
QY 901 TAGCGGTTTCACTCTACCTACGAGGCGCAGCGCAATATTTCTACGCGACTACGAGGAG 960  
DB 901 TAGCGGTTTCACTCTACCTACGAGGCGCAGCGCAATATTTCTACGCGACTACGAGGAG 960  
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTTCATCTGATACATGAGAACCTCGCCCGAGGA 1020  
DB 961 TGGCTCAACAGGATAAGCTCAAGAACCTTCATCTGATACATGAGAACCTCGCCGAGGA 1020  
QY 1021 AGCAGCGCATAGTCTACTACGATACGATGAGTCACTCTTCGTCAGGAGCGGCTACGGG 1080  
DB 1021 AGCAGCTGACATCGTCTTACTACGACACGACGAGCTGATATTCGTGAGAAACGGCTACGA 1080  
QY 1081 GACAAGCGGGGCTTATAACCTTACATCAACCTAGGCTCGAGCAAGCGCGAAGGTGGGTT 1140  
DB 1081 AGCAAGCGGGGCTGATAACATACATCAACCTTCGCTCAAGCAAGCGCGAAGGTGGGTT 1140  
QY 1141 TATGTGCCGAAGTTCGCGGGCGCTGTCATCCACGAGTATCTGTAACCTCGAGGCTGG 1200  
DB 1141 TATGTGCCGAAGTTCGCGGGCGCTGTCATCCACGAGTATCTGTAACCTCGAGGCTGG 1200  
QY 1201 GTAGACAAGTACTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
DB 1201 GTAGACAAGTACTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGAGTATGGCTTACTCCGCTGTTGGAGCTACTCGCGGGTGGGCTGA 1311  
DB 1261 GCCAAGCGGAGTATGGCTTACTCCGCTGTTGGAGCTACTCGCGGGTGGGCTGA 1311

RESULT 5  
ABX08428

ID ABX08428 standard; DNA; 1311 BP.

XX AC

XX DT

XX DE

XX KW

XX KW

XX KW

XX KW

XX KW

XX OS

XX PN

XX PD

XX PF

XX PR

XX PR

XX PR

XX PR

XX PA

XX PI

XX DR

XX DR

XX FT

XX PT

XX PS

XX SQ

21-JAN-2003 (first entry)

DNA encoding alpha amylase polypeptide #18.

Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction; starch breakdown catalysis; textile desizing; lignocellulosic fibre; enzymatic de-linking; recycled paper; high-maltose syrup; dough; high glucose syrup; corn-wet milling process; detergent; baking process; beverage; oil field; fuel ethanol; brewing process; staling; starch modification.

Unidentified.

W0200268589-A2.

06-SEP-2002.

21-FEB-2002; 2002WO-US005068.

21-FEB-2001; 2001US-0270495P.

21-FEB-2001; 2001US-0270496P.

14-MAY-2001; 2001US-0291122P.

(DIVE-) DIVERSA CORP.

Callen W, Richardson T, Frey G;

WPI: 2003-018656/01.

P-PSDB; AB030357.

Novel purified polypeptide with alpha-amylase activity, useful e.g. for liquefying starch, for textile desizing, for treating lignocellulosic fibers, and for producing high-maltose or high-glucose syrup.

Claim 1; Fig 16; 301pp; English.

The invention relates to a purified polypeptide with alpha-amylase activity and the polynucleotide encoding it. The polypeptide is useful for hydrolysing starch linkages, for catalysing the breakdown of a starch, for modifying small molecules, for liquefying starch, for washing an object, for textile desizing, for treating lignocellulosic fibers, for improving fibre properties, for enzymatic de-linking of recycled paper pulp, for producing a high-maltose or high-glucose syrup or a mixed syrup, and for increasing the flow of production fluids from a subterranean formation by removing a viscous, starch-containing, damaging fluid formed during production operations and found within the subterranean formation which surrounds a completed well bore. The polypeptide is also useful for preparing a dough or a baked product prepared from the dough and in corn-wet milling processes, detergents, baking processes, beverages, oil fields (fuel ethanol), brewing processes and starch modification in the paper and pulp industry, for removing starch containing stains from a material and for reducing staling of bakery products. Sequences ABX08411-ABX08515 represent DNA molecules encoding alpha amylase polypeptides of the invention

Sequence 1311 BP; 317 A; 373 G; 378 G; 243 T; 0 U; 0 Other;

Query Match 91.0%; Score 1192.6; DB 7; Length 1311;

Best Local Similarity 94.4%; Pred. No. 0;

Matches 1237; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1 ATGGCCAACTACTCCAGCTCGAAGAGCGGGGTCATAATCGACGGTTCCTACTGGGAC 60

DB 1 ATGGCCAACTACTCCAGCTCGAAGAGCGGGGTCATAATCGACGGTTCCTACTGGGAC 60

QY 61 GTGCCTTCAGGAGGAATATGGTGGGACACAATACGCGAAGATACCGGAGTGGTACGAT 120

DB 61 GTGCCTTCAGGAGGAATATGGTGGGACACAATACGCGAAGATACCGGAGTGGTACGAT 120

121 GCCGGAATCTCCGAATATGGATTCCCCCGGCGAGCAAGGGCATGGCGGCGCCTATTTCG 180

121 GCCGGAATCTCCGAATATGGATTCCCCCGGCGAGCAAGGGCATGGCGGCGCCTATTTCG 180

181 ATGGGCTACGACCCCTACGACTTCTTTGACCTTCGGTGAGTACGACAGAGGAACGGTA 240

181 ATGGGCTACGACCCCTACGACTTCTTTGACCTTCGGTGAGTACGACAGAGGAACGGTA 240

241 GAGACCGGCTTTGGCTCCAAAGAGGAGCTCGTGAACATGATAAACAACCCGCCCTAT 300

241 GAGACCGGCTTTGGCTCCAAAGAGGAGCTCGTGAACATGATAAACAACCCGCCCTAT 300

301 GGCATGAAGGTATACCGGATATAGTCAACACCGCGCGCGGTGACCTGGAGTGG 360

301 GGCATGAAGGTATACCGGATATAGTCAACACCGCGCGCGGTGACCTGGAGTGG 360

361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTTCGGTCAATATAC 420

361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTTCGGTCAATATAC 420

421 ACGGCCAACTACCTCGACTTCCACCCGAAAGAGCTCCATCGCGGCGATTCCCGAACATTT 480

421 ACGGCCAACTACCTCGACTTCCACCCGAAAGAGCTCCATCGCGGCGATTCCCGAACATTT 480

481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGGACAGTCTGGGCGCAGC 540

481 GGAGGCTATCCCGACATATGCCACGACAGAGCTGGGACAGTCTGGGCGCAGC 540

541 CAGGAGAGCTACCGGCGATATCTCAGGAGCATCGGCATCGATCGCTGGGCGCTTCGACTAC 600

541 CAGGAGAGCTACCGGCGATATCTCAGGAGCATCGGCATCGATCGCTGGGCGCTTCGACTAC 600

601 GTCAAGGCTATGCTCCCTGGTTCGCTCAAGGAGCTGGCTGAACCTGGTGGGAGCTGGGCG 660

601 GTCAAGGCTATGCTCCCTGGTTCGCTCAAGGAGCTGGCTGAACCTGGTGGGAGCTGGGCG 660

661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTTCTCAACTGGGCGATCTCGAGCGGT 720

661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTTTCTCAACTGGGCGATCTCGAGCGGT 720

721 GCCAAGGCTTTGACCTTCGCTCCCTCTACTACAGATGGATGAGGCTTTCACACAAACAC 780

721 GCCAAGGCTTTGACCTTCGCTCCCTCTACTACAGATGGATGAGGCTTTCACACAAACAC 780

781 ATTCCAGCGCTCGCTCTGCTCCCTTCAGAACGCGCCAGACTGTTGCTCTCCCGGACCCGTTTC 840

781 ATTCCAGCGCTCGCTCTGCTCCCTTCAGAACGCGCCAGACTGTTGCTCTCCCGGACCCGTTTC 840

841 AAGCGGTAACTTTGTAGCAACACGACGATGATATATCTGGAAACAGTATCCAGCC 900

841 AAGCGGTAACTTTGTAGCAACACGACGATGATATATCTGGAAACAGTATCCAGCC 900

901 TAGCGGTTCACTCTCACTACGAGGCGCCAGCCGACAAATATCTTACCGCACTACAGAGGAG 960

901 TAGCGGTTCACTCTCACTACGAGGCGCCAGCCGACAAATATCTTACCGCACTACAGAGGAG 960

961 TGGCTCAACAAAGGATAAGTTCGAAGAACCTCATCTGGATACATGAGAACTTCGCGCGAGGA 1020

961 TGGCTCAACAAAGGATAAGTTCGAAGAACCTCATCTGGATACATGAGAACTTCGCGCGAGGA 1020

1021 AGCACCGCATAGTCTACTACGATACGATGACTCATCTTCTCTAGGAAACCGGTACGGG 1080

1021 AGCACCGCATAGTCTACTACGATACGATGACTCATCTTCTCTAGGAAACCGGTACGGG 1080

1081 GACAAGCCGGGGCTTATACCTACATCAACCTAGGTCGAGCAAGGCCGGAAGTGGT 1140

1081 GACAAGCCGGGGCTTATACCTACATCAACCTAGGTCGAGCAAGGCCGGAAGTGGT 1140

1141 TATGTGCGGAGTTCGGGCGGCTGCATCCAGGATATCTGGTAACTCCGAGGCTGG 1200

1141 TATGTGCGGAGTTCGGGCGGCTGCATCCAGGATATCTGGTAACTCCGAGGCTGG 1200

1201 GTAGACAAAGTACCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

Dd	1201	GTGACAAAGTGGGTGGACTCAAGCGGTGGGTCTACTCGAGGCTCTGCCACGACC	1260
Qy	1261	GCCAACGGGCAGTAGTGCTACTTCGGTGTGGAGTACTCGCGGTGGGCTGA	1311
Dd	1261	GCCAAAGCGCAGTACGCGTACTTCGGTGTGGAGTACTCGCGGTGGGCTGA	1311
 RESULT 6 ABX08429			
ID	ABX08429	standard; DNA; 1311 BP.	
XX	AC	ABX08429;	
XX	AC		
DT	21-JAN-2003	(first entry)	
XX	DE		
XX	DE	DNA encoding alpha amylase polypeptide #19.	
XX	KW	Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction; starch breakdown catalysis; textile desizing; lignocellulosic fibre; enzymatic de-linking; recycled paper; high-maltose syrup; dough; high glucose syrup; corn-wet milling process; detergent; baking process; beverage; oil field; fuel ethanol; brewing process; staling;	
XX	KW	starch modification.	
XX	KW		
OS	Unidentified.		
XX	XX		
PN	WO200268589-A2.		
XX	XX		
PD	06-SEP-2002.		
XX	XX		
Pf	21-FEB-2002; 2002WO-US005068.		
XX	XX		
PR	21-FEB-2001; 2001US-0270495P.		
PR	21-FEB-2001; 2001US-0270496P.		
PR	14-MAY-2001; 2001US-0291122P.		
XX	XX		
PA	(DIVE-) DIVERSA CORP.		
XX	PI		
XX	PI	Callen W, Richardson T, Frey G;	
DR	WI:	2003-018656/01.	
DR	P-PSDB; ABU03058.		
XX	XX		
PT	Novel purified polypeptide with alpha-amylase activity, useful e.g. for liquefying starch, for textile desizing, for treating lignocellulosic fibers, and for producing high-maltose or high-glucose syrup.		
XX	XX		
XX	Claim 1; Fig 16; 30pp; English.		
CC	CC	The invention relates to a purified polypeptide with alpha-amylase activity and the polynucleotide encoding it. The polypeptide is useful for hydrolysing starch linkages, for catalysing the breakdown of a starch, for modifying small molecules, for liquefying starch, for washing an object, for textile desizing, for treating lignocellulosic fibers, for improving fibre properties, for enzymatic de-linking of recycled paper pulp, for producing a high-maltose or high-glucose syrup or a mixed syrup, and for increasing the flow of production fluids from a subterranean formation by removing a viscous, starch-containing, fluid formed during production operations and found within the subterranean formation which surrounds a completed well bore. The polypeptide is also useful for preparing a dough or a baked product prepared from the dough and in corn-wet milling processes, detergents, baking processes, beverages, oil fields (fuel ethanol), brewing processes and starch modification in the paper and pulp industry, for removing starch containing stains from a material and for reducing staling of bakery products. Sequences ABX08411-ABX08515 represent DNA molecules encoding alpha amylase polypeptides of the invention	
XX	Sequence 1311 BP; 312 A; 372 C; 380 G; 247 T; 0 U; 0 Other;		
SQ	Query Match	90.5%; Score 1186.2; DB 7; Length 1311;	
	Best Local Similarity	94.1%; Pred. No. 0;	

QY 1081 GACAAAGCGGGCTTATACCTACATCAACCTAGGCTGAGCAAGCGCGAAGTGGGTT 1140  
DB 1081 GACAAAGCGGGCTTATACCTACATCAACCTAGGCTGAGCAAGCGCGAAGTGGGTC 1140  
QY 1141 TATGTCGGAAGTTGCGGGCGGTGATCCACGAGTATACCTGAGGCTGG 1200  
DB 1141 TATGTCGGAAGTTGCGGGCGGTGATCCACGAGTATACCTGAGGCTGG 1200  
QY 1201 GTAGCAAGTACCTTACTCAAGCGGCTGGTCTATCTGCAAGCTTCAAGCTTACGACCT 1260  
DB 1201 GTAGCAAGTACCTTACTCAAGCGGCTGGTCTATCTGCAAGCTTCAAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGCTATGCTTACTCCGCTGAGCTACTCGGGGCTGGCTGA 1311  
DB 1261 GCCAAGCGGCTATGCTTACTCCGCTGAGCTACTCGGGGCTGGCTGA 1311

RESULT 7  
ABX08424  
ID ABX08424 standard; DNA; 1311 BP.  
AC  
XX ABX08424;  
DT  
XX 21-JAN-2003 (first entry)  
XX  
XX DNA encoding alpha amylase polypeptide #14.  
XX  
XX Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
XX Unidentified.  
XX  
XX WO200268589-A2.  
XX  
XX 06-SBP-2002.  
XX  
XX 21-FEB-2002; 2002WO-US005068.  
XX  
XX 21-FEB-2001; 2001US-0270495P.  
XX 21-FEB-2001; 2001US-0270496P.  
XX 14-MAY-2001; 2001US-0291122P.  
XX  
XX (DIVE-) DIVERSA CORP.  
XX  
XX Callen W, Richardson T, Frey G;  
XX  
XX WPI; 2003-018656/01.  
XX DR P-PSDB; ABU03053.  
XX  
XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
PT liquefying starch, for textile desizing, for treating lignocellulosic  
PT fibers, and for producing high-maltose or high-glucose syrup.  
XX  
XX Claim 1; Fig 16; 301pp; English.  
XX  
XX The invention relates to a purified polypeptide with alpha-amylase  
CC activity and the polynucleotide encoding it. The polypeptide is useful  
CC for hydrolysing starch linkages, for catalysing the breakdown of a  
CC starch, for modifying small molecules, for liquefying starch, for washing  
CC an object, for textile desizing, for treating lignocellulosic fibers, for  
CC improving fibre properties, for enzymatic de-inking of recycled paper  
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging  
CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC

CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of  
CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
CC encoding alpha amylase polypeptides of the invention  
XX  
SQ Sequence 1311 BP; 309 A; 367 C; 379 G; 256 T; 0 U; 0 Other;  
Query Match 90.2%; Score 1183; DB 7; Length 1311;  
Best Local Similarity 93.9%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
QY 1 ATGGCCAAAGTACTCCGAGCTGGAAAAGGGGGGGTCAATATGAGGCGTCTTACTTGGGAC 60  
DB 1 ATGGCCAAAGTACTCCGAGCTGGAAAAGGGGGGGTCAATATGAGGCGTCTTACTTGGGAC 60  
QY 61 GTGCTTTCAGGAGGAATATGTTGGGACAAATACGCGCAGAAATACCGGAGTGGTACGAT 120  
DB 61 GTGCTTTCAGGAGGAATATGTTGGGACAAATACGCGCAGAAATACCGGAGTGGGACG 120  
QY 121 GCGGGAATCTCCGAATATGATTCCCGCGGAGCAAGGGCATGCGCGCGCTATTTCG 180  
DB 121 GCGGGAATCTCCGAATATGATTCCCGCGGAGCAAGGGCATGCGCGCGCTATTTCG 180  
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGTGAGTACGACCAAGGGAACGGTA 240  
DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGTGAGTACGACCAAGGGAACGGTA 240  
QY 241 GAGACGCGCTTTCGCTCCAGCAGGAGCTCGTCAACATGATATAACACCGCCACGCTAT 300  
DB 241 GAGACGCGCTTTCGCTCCAGCAGGAGCTCGTCAACATGATATAACACCGCCACGCTAT 300  
QY 301 GGATCAAGGTATATAGCCGATATAGTCAACACCGCGCGCGGTGACTTGGAGTGG 360  
DB 301 GGATCAAGGTATATAGCCGATATAGTCAACACCGCGCGCGGTGACTTGGAGTGG 360  
QY 361 AACCCCTTCGTGAACGACTATACCTGAGCGGACTTCTCAAGGTCGGTGGTAATAC 420  
DB 361 AACCCCTTCGTGAACGACTATACCTGAGCGGACTTCTCAAGGTCGGTGGTAATAT 420  
QY 421 ACGGCCAACTACCTCGACTTCCACCGGAGCTCCATCGCGGGCGANTTCCGGAACTTT 480  
DB 421 ACGGCCAACTACCTCGACTTCCACCGGAGCTCCATCGCGGGCGANTTCCGGAACTTT 480  
QY 481 GGAGGCTATCCGACATATATCCACGAGCTGGGACCAAGTACTGGCTTGGGGCAGC 540  
DB 481 GGAGGCTATCCGACATATATCCACGAGCTGGGACCAAGTACTGGCTTGGGGCAGC 540  
QY 541 CAGGAGAGCTACGCGCATATCTCAGGAGCATCGGATCGCTGCGGCTTCGACTAC 600  
DB 541 CAGGAGAGCTACGCGCATATCTCAGGAGCATCGGATCGCTGCGGCTTCGACTAC 600  
QY 601 GTCAAGGGCTATGCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
DB 601 GTCAAGGGCTATGCTCCCTGGGTCGTCAAGGACTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
QY 661 GTTGGAGAGTACTGGGACCAACCGTCGACGCTGCTCACTGGGCATCTCAGCGGT 720  
DB 661 GTTGGAGAGTACTGGGACCAACCGTCGACGCTGCTCACTGGGCATCTCAGCGGT 720  
QY 721 GCCAAGGCTTTGACTTCGGCCCTCTACTCAAGATGGATGAGGCGCTTTGACAAACAAAC 780  
DB 721 GCCAAGGCTTTGACTTCGGCCCTCTACTCAAGATGGATGAGGCGCTTTGACAAACAAAC 780  
QY 781 ATTCCAGCGCTGCTCTGCGCTTCAGAACGGCAGACTGTTGCTCCCGCGACCCGTTTC 840  
DB 781 ATTCCAGCGCTGCTCTGCGCTTCAGAACGGCAGACTGTTGCTCCCGCGACCCGTTTC 840  
QY 841 AAGGCGCTTAACCTTTGTAGCAAAACACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
DB 841 AAGGCGCTTAACCTTTGTAGCAAAACACGACACCGATATAATCTGGAACAAAGTATCC 900  
QY 901 TAGCGGTTTCATCTCCTTACGAGGCGCGCCGACAAATATTTACCGGACTACGAGGAG 960

Db 901 TATGCTTTCATCTCACCTACGAGCCAGCCGCTCATATTCTACCGGACTACGAGGAG 960  
 QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGGAGGA 1020  
 Db 961 TGGCTCAACAGGACAGGTTGAAACCTCATATGGATACACGACCCTCGAGGGGGA 1020  
 QY 1021 AGCACCGCATAGTCTACTACGATAACGATGAATCATCTTCGTGAGGAACGGCTACGGG 1080  
 Db 1021 AGCACCGCATAGTCTACTACGATAACGATGAATCATCTTCGTGAGGAACGGCTACGGG 1080  
 QY 1081 GACAGCGGGCTTATACCTACATCACTAGCTCGACGACGAGCCGGAAGTGGTT 1140  
 Db 1081 GACAGCGGGCTTATACCTACATCACTAGCTCGACGACGAGCCGGAAGTGGTT 1140  
 QY 1141 TATGTCCGAAGTTCGCGGGCGCTGCATCCACGAGTATCTGTGAACCTCGGAGGCTGG 1200  
 Db 1141 TATGTCCGAAGTTCGCGGGCGCTGCATCCACGAGTATCTGTGAACCTCGGAGGCTGG 1200  
 QY 1201 GTAGCAAGTACGTTACTCAAGCGCTGGTCTATCTCGAAGCTCGAGTTAGACCCCT 1260  
 Db 1201 GTAGCAAGTACGTTACTCAAGCGCTGGTCTATCTCGAAGCTCGAGTTAGACCCCT 1260  
 QY 1261 GCCAACCGGCGAGTATGGCTACTCCGTGTGGAGTACTCGGGGTGGGCTGA 1311  
 Db 1261 GCCAACCGGCGAGTATGGCTACTCCGTGTGGAGTACTCGGGGTGGGCTGA 1311

## RESULT 8

ABX08416

ID ABX08416 standard; DNA; 1311 BP.

AC ABX08416;

XX

XX

DT 21-JAN-2003 (first entry)

XX

DE

XX

KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.

XX

OS Unidentified.

XX

PN WO200268589-A2.

XX

PD 06-SEP-2002.

XX

XX 21-FEB-2002; 2002WO-US005068.

PF

XX

PR 21-FEB-2001; 2001US-0270495P.

PR

PR 21-FEB-2001; 2001US-0270496P.

PR

PR 14-MAY-2001; 2001US-0291122P.

XX

XX

PA (DIVE-) DIVERSA CORP.

XX

XX

PI Callen W, Richardson T, Frey G;

XX

XX

DR P-PSDB; ABU03045.

DR

XX

XX

PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 1; Fig 16; 30pp; English.  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a

CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention  
 XX

SQ Sequence 1311 BP; 321 A; 357 C; 382 G; 251 T; 0 U; 0 Other;

## Query Match

89.4%; Score 1171.8; DB 7; Length 1311;

Best Local Similarity 93.4%; Pred. No. 0;

Matches 1224; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1 ATGCGCAAGTACTCCGAGCTTGGAAAGGGGGGTCATATGCGAGCGTTCTACTCGGAC 60

Db 1 ATGCGCAAGTACTCCGAGCTTGGAAAGGGGGGTCATATGCGAGCGTTCTACTCGGAC 60

QY 61 GTCCCTTACGAGGAATATGTTGGGACACAATACGGCAGAAATACCGAGTGGTACGAT 120

Db 61 GTCCCTTACGAGGAATATGTTGGGACACAATACGGCAGAAATACCGAGTGGTACGAT 120

QY 121 GCCGGATCTCCGCAATATGGATTCGCCGGGAGCAAGGCGATGGCGCGCTATTTCG 180

Db 121 GCCGGATCTCCGCAATATGGATTCGCCGGGAGCAAGGCGATGGCGCGCTATTTCG 180

QY 181 ATGGGCTACGACCCCTACGACTTCCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTA 240

Db 181 ATGGGCTACGACCCCTACGACTTCCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTA 240

QY 241 GAGACGCGTTTGGCTTCCAGAGGAGCTCGTGAACATGATATAACACCGCCACGCTAT 300

Db 241 GAGACGCGTTTGGCTTCCAGAGGAGCTCGTGAACATGATATAACACCGCCACGCTAT 300

QY 301 GGCATGAAGTATATAGCCATATAGTCATCAACACCGCGCGGTGACCTCGGAGTGG 360

Db 301 GGCATGAAGTATATAGCCATATAGTCATCAACACCGCGCGGTGACCTCGGAGTGG 360

QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCCTCAAGGTTCGCGTAAATAC 420

Db 361 AACCCCTTCGTGGGACTACACCTGGACCGACTTCCTCAAGGTTCGCGTAAATAT 420

QY 421 ACGGCCAAGTACTCGACTTCACCCGACGAGCTCCATGCGGGCGATTCCGGAACATTT 480

Db 421 ACTGCCAAGTACTCGACTTCACCCGACGAGCTCCATGCGGGCGATTCCGGAACATTT 480

QY 481 GAGAGGCTATCCCGACATATGCGCACCAAGAGCTGGGACCACTGCGTCTGGGCGAGC 540

Db 481 GAGAGGCTATCCCGACATATGCGCACCAAGAGCTGGGACCACTGCGTCTGGGCGAGC 540

QY 541 CAGGAGCTACCGGCGATATCTCAGGAGCATCGGATCGATGCTGGCGCTTCGACTAC 600

Db 541 GATGAGAGCTACCGCGCTACCTAAGGAGCATGGCGCTTGATGCTGGCGCTTCGACTAC 600

QY 601 GTCAGGGCTATGCTCCCTGGGTGCTCAAGGACTGGCTCAACTGGTGGGAGGCTGGGG 660

Db 601 GTCAGGGCTATGAGGAGCTGGGTGCTCAAGGACTGGCTCAACTGGTGGGAGGCTGGGG 660

QY 661 GTTGGAGAGTACTGGGACACCAACGTCGAGCGCTGTTCTCAACTGGGCATACCTGAGCGGT 720

Db 661 GTCGGGAGTACTGGGACACCAACGTTGATGCACTGCTCAACTGGGCCTACTCGAGCGAT 720

QY 721 GCCAAGGCTTTTGACTTCGCCCTCTACTACAGATGATGAGCGCTTTCGACACAAAAC 780

Db 721 GCCAAGGCTTTTGACTTCGCCCTCTACTACAGATGATGAGCGCTTTCGACACAAAAC 780

QY 781 ATTCCAGCGCTGCTCTCTCCCTTCCAGACGGCCAGACTGTTGTCTCCGCGACCCGTTCC 840  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 781 ATTCCAGCGCTGCTCTCTCCCTTCCAGACGGCCAGACTGTTGTCTCCGCGACCCGTTCC 840  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 841 AAGCGCGTAACCTTTGTAGCAACACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 841 AAGCGCGTAACCTTTGTAGCAACACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 901 TAGCGCTTATCTCTACCTACGAGGGCCAGCCGACATATTCTACCGGACTACGAGGAG 960  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 901 TAGCGCTTATCTCTACCTACGAGGGCCAGCCGACATATTCTACCGGACTACGAGGAG 960  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 961 TGCGCTCAACAAGGATAGCTCAAGAACCTCATCTGATATCATGAGAACCTCCCGGAGGA 1020  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 961 TGCGCTCAACAAGGATAGCTCAAGAACCTCATCTGATATCATGAGAACCTCCCGGAGGA 1020  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1021 AGACCGCATAGTCTACTACGATACGATGATCACTCTGTCAGGAAACGGCTACGGG 1080  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1021 AGACCTGACATAGTCTACTACGATACGATGATCACTCTGTCAGGAAACGGCTACGGG 1080  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1081 GACAACCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCCGAGAGTGGGT 1140  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1081 GACAACCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCCGAGAGTGGGT 1140  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1141 TATGTCGCAAGTTCGCGGGCGGTGCATCCACGAGTATCTGGTAACTCGGAGGCTGG 1200  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1141 TATGTCGCAAGTTCGCGGGCGGTGCATCCACGAGTATCTGGTAACTCGGAGGCTGG 1200  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1201 GTAGACAGTACGCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1201 GTAGACAGTACGCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1261 GCCAACGGCAGTATGGCTACTCGGTGGAGCTACTGCGGGTGGGCTGA 1311  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 QY 1261 GCCAACGGCAGTATGGCTACTCGGTGGAGCTACTGCGGGTGGGCTGA 1311  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

## RESULT 9

ABX08440

ID ABX08440 standard; DNA; 1311 BP.

AC ABX08440;

XX

XX

DT 21-JAN-2003 (first entry)

XX

XX

DE DNA encoding alpha amylase polypeptide #30.

XX

XX

KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough; baking process;  
 KW high glucose syrup; corn-wet milling process; detergent; brewing process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.

XX

OS Unidentified.

XX

XX

PN W020268589-A2.

XX

XX

PD 06-SEP-2002.

XX

XX

XX 21-FEB-2002; 2002WO-US005068.

XX

XX

PR 21-FEB-2001; 2001US-0270495P.

XX

XX

PR 21-FEB-2001; 2001US-0270496P.

XX

PR 14-MAY-2001; 2001US-0291122P.

XX

XX

(DIVE-) DIVERSA CORP.

PA

XX

XX

Callen W, Richardson T, Frey G;

PI

XX

XX

WPI; 2003-018656/01.

DR

XX

P-PSDB; ABU03069.

DR

XX

PT

Novel purified polypeptide with alpha-amylase activity, useful e.g. for

PT

liquefying starch, for textile desizing, for treating lignocellulosic

PT

fibers, and for producing high-maltose or high-glucose syrup.

XX

FS

Claim 1; Fig 16; 301pp; English.

XX

XX

The invention relates to a purified polypeptide with alpha-amylase

CC

activity and the polynucleotide encoding it. The polypeptide is useful

CC

for hydrolysing starch linkages, for catalysing the breakdown of a

CC

starch, for modifying small molecules, for liquefying starch, for washing

CC

an object, for textile desizing, for treating lignocellulosic fibers, for

CC

improving fibre properties, for enzymatic de-inking of recycled paper;

CC

pulp, for producing a high-maltose or high-glucose syrup or a mixed

CC

symp, and for increasing the flow of production fluids from a

CC

subterranean formation by removing a viscous, starch-containing,

CC

fluid formed during production operations and found within the

CC

subterranean formation which surrounds a completed well bore. The

CC

polypeptide is also useful for preparing a dough or a baked product

CC

prepared from the dough and in corn-wet milling processes, detergents,

CC

baking processes, beverages, oil fields (fuel ethanol), brewing processes

CC

and starch modification in the paper and pulp industry, for removing

CC

starch containing stains from a material and for reducing staling of

CC

bakery products. Sequences ABX08411-ABX08515 represent DNA molecules

CC

encoding alpha amylase polypeptides of the invention

XX

SQ

Sequence 1311 BP; 324 A; 367 C; 377 G; 243 T; 0 U; 0 Other;

Query Match

88.5%; Score 1160.6; DB 7; Length 1311;

Best Local Similarity

92.88; Pred. No. 0;

Matches 1217; Conservative

0; Mismatches 94; Indels 0; Gaps 0;

QY

1

ATGSCCAAGTACTCCGAGCTGGAAAGGGGGGGTATAATGCAAGGCTTCTACTGGAC

60

Db

1

ATGSCCAAGTACTCCGAGCTGGAAAGGGGGGGTATAATGCAAGGCTTCTACTGGAC

60

QY

61

GTGCTCTCAGGAGGATATGTTGGGACACAATACGGCAGAGATACCGGAGTGGTACGAT

120

Db

61

GTGCTCTCAGGAGGATATGTTGGGACACAATACGGCAGAGATACCGGAGTGGTACGAG

120

QY

121

GCGGAAATCTCCGCAATATGATTTCCCGCGCAGCAAGGCGATGCGGCGGCTATTGCG

180

Db

121

GCGGAAATATCCGCCATTTGGAATTCCTCCCGCAGCAAGGCTATGAGCGCGCTATTGCG

180

QY

181

ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTA

240

Db

181

ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTG

240

QY

241

GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACACCGCCACGCTAT

300

Db

241

GAAACGAGGTTCCGCTCAAAGCAGGAGCTCATAAACATGATAAACACCGCCACGCTAC

300

QY

301

GGCATGAAGTATAGCCGATATAGTATCAACACCGCGCGCGGCTGACCTGGAGTGG

360

Db

301

GGCATGAAGTATAGCCGATATAGTATCAACACCGCGCGCGGCTGACCTGGAGTGG

360

QY

361

AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTTCAAAGGTCGCGTGGTAAATAC

420

Db

361

AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTTCAAAGGTCGCGTGGTAAATAC

420

QY

421

ACGGCCAACTACTCTGACTTCCACCGAAGAGTCCATCGGCGGATTCGGGAACATTT

480

Db

421

ACGGCCAACTACTCTGACTTCCACCGAAGAGTCCATCGGCGGATTCGGGAACATTT

480

QY

481

GGAGGCTATCCGACATATGCCACGAGCTGGGACCAAGTACTGGCTCTCGGCGCAGC

540

Db

481

GGAGGCTATCCGACATATGCCACGAGCTGGGACCAAGTACTGGCTCTCGGCGCAGC

540

QY

541

CAGGAGAGTACCGGATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAC

600

Db

541

CAGGAGAGTACCGGATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAC

600

QY

601

GTCAAGGGCTATGCTCCCTGGGTCGTCAAGACTGGGCTGAACCTGGTGGGAGGCTGGCGG

660



Db 601 GTCAGGCTACGAGCTGGTCTGTCAGAGACTGGCTGGAGGCTGGGCC 660  
QY 661 GTTGAGAGTACTGGGACCAACGTCGACGCTGTTCTCAATGGGCATATCGAGCGT 720  
Db 661 GTCGGGAGTACTGGGACCAACGTTGATGACTGCTCAATGGGCCTACTCGAGCGAT 720  
QY 721 GCCAAGGCTTTGACTTCGCTCGCTCTACTACAGATGGATAGGCTTTGACAACAAAC 780  
Db 721 GCAAAGTCTTCGACTTCGCTCTACTACAGATGGATAGGCTTTGACAACAAAC 780  
QY 781 ATTCAGCGCTGCTCTGCTCTGCTCTCAGAACGCCAGACTGTTGTCTCCCGGACCCGTTTC 840  
Db 781 ATTCAGCGCTGCTCTGCTCTGCTCTCAGAACGCCAGACTGTTGTCTCCCGGACCCGTTTC 840  
QY 841 AAGCGGTAACTTTGTTAGCAACACGACACCGATTAATCTGGAAAGTATCGAGCC 900  
Db 841 AAGCGGTAACTTTGTTAGCAACACGATGACACCGATTAATCTGGAAAGTATCGAGCC 900  
QY 901 TACGGTTTCATCTCCTCAGAGGCGCAGCGCAATATTCTACCGGACTACGAGAG 960  
Db 901 TACGGTTTCATCTCCTCAGAGGCGCAGCGCAATATTCTACCGGACTACGAGAG 960  
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGGAGGA 1020  
Db 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCGGAGGA 1020  
QY 1021 AGCAGCGATAGTCTACTAGATAAGATGAATCATCTTCGTCAGGAAGGCTACGGG 1080  
Db 1021 AGCAGCGATAGTCTACTAGATAAGATGAATCATCTTCGTCAGGAAGGCTACGGG 1080  
QY 1081 GACAGCGGGGCTTAACTCACTACATCACTAGGCTCGAGAGCGGAGAGTGGGTT 1140  
Db 1081 GACAGCGGGGCTTAACTCACTACATCACTAGGCTCGAGAGCGGAGAGTGGGTT 1140  
QY 1141 TATGTGCGGAGTTCGCGGCGCTGCTGATCCACAGATATCTGTAACTCGAGGCTGG 1200  
Db 1141 TATGTGCGGAGTTCGCGGCGCTGCTGATCCACAGATATCTGTAACTCGAGGCTGG 1200  
QY 1201 GTAGCAAGTACGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Db 1201 GTAGCAAGTACGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGAGTATGGTACTTCGCTGTGAGACTACTCGCGGGTGGGCTGA 1311  
Db 1261 GCCAAGCGGAGTATGGTACTTCGCTGTGAGACTACTCGCGGGTGGGCTGA 1311

RESULT 10  
ABX08437

ID ABX08437 standard; DNA; 1311 BP.

AC ABX08437;

XX

DT 21-JAN-2003 . (first entry)

XX

DE DNA encoding alpha amylase polypeptide #27.

XX

KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.

XX

OS Unidentified.

XX

FN WO200268599-A2.

XX

PD 06-SEP-2002.

XX

PF 21-FEB-2002; 2002WO-US005068.

XX

PR 21-FEB-2001; 2001US-0270495P.  
PR 21-FEB-2001; 2001US-0270496P.  
XX 14-MAY-2001; 2001US-0291122P.  
PA (DIVE-) DIVERSA CORP.  
XX

PI Callen W, Richardson T, Frey G;

XX WPI; 2003-018656/01.

DR P-PSDB; ABU03066.

XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
PT liquefying starch, for textile desizing, for treating lignocellulosic  
PT fibers, and for producing high-maltose or high-glucose syrup.  
XX

PS Claim 1; Fig 16; 30pp; English.

XX The invention relates to a purified polypeptide with alpha-amylase  
CC activity and the polynucleotide encoding it. The polypeptide is useful  
CC for hydrolysing starch linkages, for catalysing the breakdown of a  
CC starch, for modifying small molecules, for liquefying starch, for washing  
CC an object, for textile desizing, for treating lignocellulosic fibers, for  
CC improving fibre properties, for enzymatic de-inking of recycled paper  
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging  
CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of  
CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
CC encoding alpha amylase polypeptides of the invention  
XX

SQ Sequence 1311 BP; 308 A; 368 C; 379 G; 256 T; 0 U; 0 Other;

Query Match

Best Local Similarity 88.2%; Score 1155.8; DB 7; Length 1311;

Matches 1214; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1 ATGGCCAACTACTCCAGCTCGAAGAGGCGGGGCTCATATGACGGCTTCTACTGGGAC 60  
Db 1 ATGGCCAACTACTCCAGCTCGAAGAGGCGGGGCTCATATGACGGCTTCTACTGGGAC 60  
QY 61 GTGCTTCAGGAGGATATGTTGGGACACAATACGCGAAGATACCGAGTGGTACGAT 120  
Db 61 GTGCTTCAGGAGGATATGTTGGGACACAATACGCGAAGATACCGAGTGGTACGAT 120  
QY 121 GCGGGAATCTCCGCAATATGGAATTCCTCCCGGAGCAAGGGCATGGCGGCGCTTATCG 180  
Db 121 GCGGGAATCTCCGCAATATGGAATTCCTCCCGGAGCAAGGGCATGGCGGCGCTTATCG 180  
QY 181 ATGGGCTACGCCCTTACGACTTCTTGACCTCGGTGAGTACGACGAGGACGGTA 240  
Db 181 ATGGGCTACGCCCTTACGACTTCTTGACCTCGGTGAGTACGACGAGGACGGTA 240  
QY 241 GAGACGCGCTTTGGTCTCCAAAGCAGAGCTCTGCTGAACATGATAAACCACCGCCTAT 300  
Db 241 GAGACGCGCTTTGGTCTCCAAAGCAGAGCTCTGCTGAACATGATAAACCACCGCCTAT 300  
QY 301 GGCATGAAGGTAATAGCCGATATAGTATCAACACCGCGCGGGGCTGACCTGAGTGG 360  
Db 301 GGCATGAAGGTAATAGCCGATATAGTATCAACACCGCGCGGGGCTGACCTGAGTGG 360  
QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTCGCGGTAAATAC 420  
Db 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAGGTCGCGGTAAATAC 420  
QY 421 ACGGCCAACTACTCGACTTCGACTTCCACCGAACGAGCTCCATGCGGGCGATTCGGAACATTT 480  
Db 421 ACGGCCAACTACTCGACTTCGACTTCCACCGAACGAGCTCCATGCGGGCGATTCGGAACATTT 480

QY 481 GGAGGCTATCCCGACATATGCGACAGAGCTGGACCAAGTACTGCTCTGGGCCAGC 540  
DB 481 CGAGGCTATCCCGACATATGCGACAGAGCTGGACCAAGTACTGCTCTGGGCCAGC 540  
QY 541 CAGGAGCTAGCGGCTATCTCAGGAGCATCGGATCGATCGCTGGGCTTGGACTAC 600  
DB 541 CAGGAGCTAGCGGCTATCTCAGGAGCATCGGATCGATCGCTGGGCTTGGACTAC 600  
QY 601 GTCAAGGCTATGCTCCCTGGGCTCGTCAAGAGTCTGGTGAATCGTGGGAGGCTGGGCG 660  
DB 601 GTCAAGGCTATGCTCCCTGGGCTCGTCAAGAGTCTGGTGAATCGTGGGAGGCTGGGCG 660  
QY 661 GTTGGAGTACTGGGACACCAAGTCTGAGCGTCTTCAAGTGGGATCTCGAGCGGT 720  
DB 661 GTTGGAGTACTGGGACACCAAGTCTGAGCGTCTTCAAGTGGGATCTCGAGCGGT 720  
QY 721 GCCAAGGCTTTGACTTTCGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAAC 780  
DB 721 GCCAAGGCTTTGACTTTCGCTCTACTACAAGATGGATGAGGCTTTGACAAACAAAC 780  
QY 781 ATTCCAGGCTGCTCTCTGCTTCAAGACCGGACAGTGTGTCTCCGCGACCGCTTC 840  
DB 781 ATTCCAGGCTGCTCTCTGCTTCAAGACCGGACAGTGTGTCTCCGCGACCGCTTC 840  
QY 841 AAGGCGTAACTTTGACCAACCAACGACCGATATATCTGGAAACAAGTATCCAGCC 900  
DB 841 AAGGCGTAACTTTGACCAACCAACGACCGATATATCTGGAAACAAGTATCCAGCC 900  
QY 901 TAGCGCTTATCTCTACCTACGAGGCGGACGACATATCTCCGCGACTACGAGAG 960  
DB 901 TAGCGCTTATCTCTACCTACGAGGCGGACGACATATCTCCGCGACTACGAGAG 960  
QY 961 TGGCTCAACAGGATAGCTCAGACCTCATCTGATATGATGATGATGATGATGATG 1020  
DB 961 TGGCTCAACAGGATAGCTCAGACCTCATCTGATATGATGATGATGATGATGATG 1020  
QY 1021 AGCAGCGATAGCTTACTACGATACGATGATGATGATGATGATGATGATGATG 1080  
DB 1021 AGCAGCGATAGCTTACTACGATACGATGATGATGATGATGATGATGATGATG 1080  
QY 1081 GACAGCGGCTTATATCTATCTACATCACTAGCTAGCTCAGACGCGGAGGCTT 1140  
DB 1081 AGCAAGCTGGCTTATATCTATCTATCACTCACTCGCTCAGACGAGGCTT 1140  
QY 1141 TATGTCGAGGCTTCCGCGGCTGCTATCCAGAGTACTGATGATGATGATGATGATG 1200  
DB 1141 TATGTCGAGGCTTCCGCGGCTGCTATCCAGAGTACTGATGATGATGATGATGATG 1200  
QY 1201 GTAGCAAGTACTTACTCAAGCGCTGGTCTATCTCGAAGCTTCCAGCTTACGACCT 1260  
DB 1201 GTAGCAAGTACTTACTCAAGCGCTGGTCTATCTCGAAGCTTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGCTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1311  
DB 1261 GCCAAGCGGCTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1311

RESULT 11  
ID ABX08422  
XX ABX08422 standard; DNA; 1311 BP.  
AC ABX08422;  
DT 21-JAN-2003 (first entry)  
XX DNA encoding alpha amylase polypeptide #12.  
DE Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
XX starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;

starch modification.  
Unidentified.  
W0200268589-A2.  
06-SEP-2002.  
21-FEB-2002; 2002WO-US005068.  
21-FEB-2001; 2001US-0270495P.  
21-FEB-2001; 2001US-0270496P.  
14-MAY-2001; 2001US-0291122P.  
(DIVE-) DIVERSA CORP.  
Callen W, Richardson T, Frey G;  
WPI; 2003-018656/01.  
P-PSDB; ABU03051.  
Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
liquefying starch, for textile desizing, for treating lignocellulosic  
fibers, and for producing high-maltose or high-glucose syrup.  
Claim 1; Fig 16; 301pp; English.  
The invention relates to a purified polypeptide with alpha-amylase  
activity and the polynucleotide encoding it. The polypeptide is useful  
for hydrolysing starch linkages, for catalysing the breakdown of a  
starch, for modifying small molecules, for liquefying starch, for washing  
an object, for textile desizing, for treating lignocellulosic fibers, for  
improving fibre properties, for enzymatic de-inking of recycled paper  
pulp, for producing a high-maltose or high-glucose syrup or a mixed  
syrup, and for increasing the flow of production fluids from a  
subterranean formation by removing a viscous, starch-containing, damaging  
fluid formed during production operations and found within the  
subterranean formation which surrounds a completed well bore. The  
polypeptide is also useful for preparing a dough or a baked product  
prepared from the dough and in corn-wet milling processes, detergents,  
baking processes, beverages, oil fields (fuel ethanol), brewing processes  
and starch modification in the paper and pulp industry, for removing  
starch containing stains from a material and for reducing staling of  
bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
encoding alpha amylase polypeptides of the invention  
Sequence 1311 BP; 313 A; 369 C; 381 G; 248 T; 0 U; 0 Other;

Query Match 87.68; Score 1147.8; DB 7; Length 1311;  
Best Local Similarity 92.28; Pred. No. 3e-312;  
Matches 1209; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
QY 1 ATGCCCAAGTACTCCGAGCTCGAAAGGCGGGGTCTAATGACGGCTTCTACTGGAC 60  
DB 1 ATGCCCAAGTACTCCGAGCTCGAAAGGCGGGGTCTAATGACGGCTTCTACTGGAC 60  
QY 61 GTGCTTCAGAGGATATGCTGGACACATACGGCAGAGATACCGGAGTGTAGCAT 120  
DB 61 GTCCCAATGGAGGAATCTGGTGGACACATAGCCAGGAATATCCGACTGGGACG 120  
QY 121 GCCCGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGGCATGGCGCGCTATTTCG 180  
DB 121 GCCCGAATCTCCGCGGATATGATTTCTCCCGGAGCAAGGGGTATGAGCGCGCTATTTCG 180  
QY 181 ATGGGCTACAGCCCTTACGACTCTTTGACTCGGTGAGTACGACAGAGGGAACGGTA 240  
DB 181 ATGGGCTACAGCCCTTACGACTCTTTGACTCGGTGAGTACGACAGAGGGAACGGTG 240  
QY 241 GAGACGCGCTTTGGCTCCAGAGGAGCTCGTGAACATGATAAACCAGCCCGCCCTAT 300  
DB 241 GAAACAGGTTTCGCTCAAGAGGAGCTCTAATAATGATNAACACCGCCCGCTAT 300  
QY 301 GGCATGAAGTATATAGCCGATATAGTATCATCAACACCGCGCGGTGACCTGGAGTG 360

DB 301 GGCATGAAGGTAAATAGCCGATATAGTATCAACCAACCGCGCGGTGACCTGGAGTGG 360  
QY 361 AACCCCTTCGTGAAAGACTATACCTGGACCGGACTTCTCAAAAGGTGCGTCCGGTAAATAC 420  
DB 361 AACCCCTTCGTGAAAGACTATACCTGGACCGGACTTCTCAAAAGGTGCGTCCGGTAAATAC 420  
QY 421 ACGGCCAACTACCTCGACTTCCACCCGACAGCTCCATGCGGGCGAATCCGGGAACATTT 480  
DB 421 ACGGCCAACTACCTCGACTTCCACCCGACAGCTCCATGCGGGCGAATTCGGGAACATTT 480  
QY 481 GGAGGCTATCCCGACATATCCACAGAGCTGGGACCGATCTGGCTTCGGCCAGC 540  
DB 481 GGAGGCTATCCCGACATATCCACAGAGCTGGGACCGATCTGGCTTCGGCCAGC 540  
QY 541 CAGGAGACTACCGGCGCATATCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGACTAC 600  
DB 541 CAGGAGACTACCGGCGCATATCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGACTAC 600  
QY 601 GTCAGGCGTATGCTCCCTCGGTGCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGG 660  
DB 601 GTCAGGCGTATGCTCCCTCGGTGCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGCGG 660  
QY 661 GTTGAGAGTACTGGGACACCAAGCTCGACGCTGTTCTCAACTGGGATATCTCGAGCGGT 720  
DB 661 GTTGAGAGTACTGGGACACCAAGCTCGACGCTGTTCTCAACTGGGATATCTCGAGCGGT 720  
QY 721 GCCAAGGCTTTGACTTCGCGCTCTACTACAGATGGATGAGGCTTTGACAAACAAAC 780  
DB 721 GCCAAGGCTTTGACTTCGCGCTCTACTACAGATGGATGAGGCTTTGACAAACAAAC 780  
QY 781 ATCCAGCGCTCGCTCTCGCGCTTCAGAACCGGACGAGCTGTTGCTCCGCGACCGCTTC 840  
DB 781 ATCCAGCGCTCGCTCTCGCGCTTCAGAACCGGACGAGCTGTTGCTCCGCGACCGCTTC 840  
QY 841 AAGCGCGTAACCTTTGTAGCAAAACACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900  
DB 841 AAGCGTGAGCTTTGTAGCAAAACACGATACCGATATAATCTGGAACAAAGTATCCAGCC 900  
QY 901 TACGCGTTCATCTCACTACGAGGCGGACCGGACATATTTCTACCGGACTACGAGGAG 960  
DB 901 TACGCGTTCATCTCACTACGAGGCGGACCGGACATATTTCTACCGGACTACGAGGAG 960  
QY 961 TGGCTCAACAGGATAGCTCAAGAACTCTATGATATACATGAGAACCTTCGCGGAGGA 1020  
DB 961 TGGCTCAACAGGATAGCTCAAGAACTCTATGATATACATGAGAACCTTCGCGGAGGA 1020  
QY 1021 AGCACCGCATAGTCTACTACGATAACGATGAATCATCTTCGTCAGGAACGGCTACGGG 1080  
DB 1021 AGCACCGCATAGTCTACTACGATAACGATGAATCATCTTCGTCAGGAACGGCTACGGG 1080  
QY 1081 GACAAGCGGGGCTTATAACCTCATCACTAGGCTCGAGCAGCGGAGAGTGGGTT 1140  
DB 1081 GACAAGCGGGGCTTATAACCTCATCACTAGGCTCGAGCAGCGGAGAGTGGGTT 1140  
QY 1141 TATGTGCCAAGTTCGGGGCGGTGATCCACAGATATCTGTTAACTTCGAGGCTGG 1200  
DB 1141 TACGTTCCGAAGTTCGCGAGGCTCGTGATACACGAGTACACCGGCAATCTCGCGGCTGG 1200  
QY 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGTCTATCTCGAAGCTTCAGCTTACGACCT 1260  
DB 1201 GTGACAAGTGGGTTGAGCTCAAGCGGCTGGTCTACCTCGAGGCTCTGCTGCCACGACCCG 1260  
QY 1261 GCCAACCGGCAAGTATGGCTACTCCGCTGTGAGCTACTGCGGGTGGGCTGA 1311  
DB 1261 GCCAACCGGCAAGTATGGCTACTCCGCTGTGAGCTATTTGCGGTGTTGGTGA 1311

RESULT 12  
ABX08432  
ID ABX08432 standard; DNA; 1311 BP.  
XX  
AC ABX08432;

XX 21-JAN-2003 (first entry)  
XX  
XX DNA encoding alpha amylase polypeptide #22.  
XX  
KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
OS Unidentified.  
XX  
XX WO200268589-A2.  
XX  
XX 06-SEP-2002.  
XX  
XX 21-FEB-2002; 2002WO-US005068.  
XX  
XX 21-FEB-2001; 2001US-0270495P.  
XX  
XX 21-FEB-2001; 2001US-0270496P.  
XX  
XX 14-MAY-2001; 2001US-0291122P.  
XX  
XX (DIVE-) DIVERSA CORP.  
XX  
XX Callen W, Richardson T, Frey G;  
XX  
XX WPI; 2003-018656/01.  
XX  
XX P-PSDB; ABU03061.  
XX  
XX Novel purified polypeptide: with alpha-amylase activity, useful e.g. for  
XX liquefying starch, for textile desizing, for treating lignocellulosic  
XX fibers, and for producing high-maltose or high-glucose syrup.  
XX  
XX Claim 1; Fig 16; 301pp; English.  
XX  
XX The invention relates to a purified polypeptide with alpha-amylase  
XX activity and the polynucleotide encoding it. The polypeptide is useful  
XX for hydrolysing starch linkages, for catalysing the breakdown of a  
XX starch, for modifying small molecules, for liquefying starch, for washing  
XX an object, for textile desizing, for treating lignocellulosic fibers, for  
XX improving fibre properties, for enzymatic de-inking of recycled paper,  
XX pulp, for producing a high-maltose or high-glucose syrup or a mixed  
XX syrup, and for increasing the flow of production fluids from a  
XX subterranean formation by removing a viscous, starch-containing,  
XX fluid formed during production operations and found within the  
XX subterranean formation which surrounds a completed well bore. The  
XX polypeptide is also useful for preparing a dough or a baked product  
XX prepared from the dough and in corn-wet milling processes, detergents,  
XX baking processes, beverages, oil fields (fuel ethanol), brewing processes  
XX and starch modification in the paper and pulp industry, for removing  
XX starch containing stains from a material and for reducing staling of  
XX bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
XX encoding alpha amylase polypeptides of the invention  
XX  
XX Sequence 1311 BP; 313 A; 365 C; 382 G; 251 T; 0 U; 0 Other;

Query Match 87.6%; Score 1147.8; DB 7; Length 1311;  
Best Local Similarity 92.2%; Pred. No. 3e-312;  
Matches 1209; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
QY 1 ATGGCCCAAGTACTCCGAGCTGGAAAGGGCGGGTCAATAATCGACGGCTTCTACTGGGAC 60  
DB 1 ATGGCCCAAGTACTCCGAGCTGGAAAGGGCGGGTCAATAATCGACGGCTTCTACTGGGAC 60  
QY 61 GTGCCAGGTGGAGGAATCTGGTGGGACACCAATACGAGCAAGATACCGGAGTGGTACCAT 120  
DB 61 GTGCCAGGTGGAGGAATCTGGTGGGACACCAATACGAGCAAGATACCGGAGTGGTACCAT 120  
QY 121 GCGGGAATCTCCGCAATATGATTTCCCGGCGACAGGGCATGGCGGCGCTTATTCG 180  
DB 121 GCGGGAATATCGCCATTGGATTCCCGGCGGAGCAAGGGCATGGCGGCGCTTATTCG 180

QY 181 ATGGCTAGGACCTTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTA 240  
 Db 181 ATGGCTAGGACCTTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGGAACGGTA 240  
 QY 241 GAGAGCGCTTTGGCTTCAAGCAGAGAGCTCGTGAACATGATAAACAACCCGCCACGCTTAT 300  
 Db 241 GAGAGCGCTTTGGCTTCAAGCAGAGAGCTCGTGAACATGATAAACAACCCGCCACGCTTAT 300  
 QY 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACACCGCGCGCGGTGACTCGAGTGG 360  
 Db 301 GGCATGAAGGTAATAGCCGATATAGTCAATCAACACCGCGCGCGGTGACTCGAGTGG 360  
 QY 361 AACCCTTCTGTAAGACTATACCTGAGCGGCTTCTCAAGAGTCCGCTCGGTGAATAATAC 420  
 Db 361 AACCCTTCTGTAAGACTATACCTGAGCGGCTTCTCAAGAGTCCGCTCGGTGAATAATAC 420  
 QY 421 ACGGCAACTACTCGACTTCCACCGAAGAGCTCCATGCGGGCGATTCGGAACATTT 480  
 Db 421 ACGGCAACTACTCGACTTCCACCGAAGAGCTCCATGCGGGCGATTCGGAACATTT 480  
 QY 481 GAGGCTATCCGACATATGCCACGACAGAGCTGCGCTCTGGGCCAGC 540  
 Db 481 GAGGCTATCCGACATATGCCACGACAGAGCTGCGCTCTGGGCCAGC 540  
 QY 541 CAGGAGGCTACGCGGCTATCTCAGGAGCATCGCATCGATGCTCGGCTTCGACTTAC 600  
 Db 541 CAGGAGGCTACGCGGCTATCTCAGGAGCATCGCATCGATGCTCGGCTTCGACTTAC 600  
 QY 601 GTCAAGGCTATGCTCCCTGGTCTCAAGAGCTGGCTGACTGGTGGGAGGCTGGCG 660  
 Db 601 GTCAAGGCTATGCTCCCTGGTCTCAAGAGCTGGCTGACTGGTGGGAGGCTGGCG 660  
 QY 661 GTTGGAGAGTACTGGGACACCAACGAGCTGCTTCTCAACTGGGCTACTCGAGCGGT 720  
 Db 661 GTTGGAGAGTACTGGGACACCAACGAGCTGCTTCTCAACTGGGCTACTCGAGCGGT 720  
 QY 721 GCAAGGCTTTGACTTGGCTTCTACTCAAGAGTGGATGAGGCTTTGACAAACAAAC 780  
 Db 721 GCAAGGCTTTGACTTGGCTTCTACTCAAGAGTGGATGAGGCTTTGACAAACAAAC 780  
 QY 781 ATTCCAGGCTGCTCTGCTTCAAGAGCTGCTTCTCGCGAGCTGCTTCTCGCGAGCTTC 840  
 Db 781 ATTCCAGGCTGCTCTGCTTCAAGAGCTGCTTCTCGCGAGCTGCTTCTCGCGAGCTTC 840  
 QY 841 AAGGCGGTAACTTTGTAGCAACACGACACCGATATATCTGGAACAAAGTATCCAGCC 900  
 Db 841 AAGGCGGTAACTTTGTAGCAACACGACACCGATATATCTGGAACAAAGTATCCAGCC 900  
 QY 901 TAGCGTTTCTCTACCTACGAGGCGGCGGACGATATCTACCGGACTACGAGAG 960  
 Db 901 TAGCGTTTCTCTACCTACGAGGCGGCGGACGATATCTACCGGACTACGAGAG 960  
 QY 961 TGCGTCAACAGGATAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGA 1020  
 Db 961 TGCGTCAACAGGATAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGA 1020  
 QY 1021 AGCAGGACATAGTCTACTAGTAAACGATGAAGTCACTTCTGTCAGGAACGCTACGG 1080  
 Db 1021 AGCAGGACATAGTCTACTAGTAAACGATGAAGTCACTTCTGTCAGGAACGCTACGG 1080  
 QY 1081 GACAAGCGGGGTTTAACTCACTCACTAGGCTCGAGCAAGCGGCGAAGGTTGGT 1140  
 Db 1081 GACAAGCGGGGTTTAACTCACTCACTAGGCTCGAGCAAGCGGCGAAGGTTGGT 1140  
 QY 1141 TATGTGCGGAGTTGCGGGCGGTGATCCACGAGTATCTGTTAACTTCGAGGCTGG 1200  
 Db 1141 TATGTGCGGAGTTGCGGGCGGTGATCCACGAGTATCTGTTAACTTCGAGGCTGG 1200  
 QY 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGTCTTCTGAGCTCCAGTTTACGACCT 1260  
 Db 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGTCTTCTGAGCTCCAGTTTACGACCT 1260

QY 1261 GCCAACGGCAGTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGCTGA 1311  
 Db 1261 GCCAACGGCAGTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGCTGA 1311  
 RESULT 13  
 ABX08412  
 ID ABX08412 standard; DNA; 1311 BP.  
 AC ABX08412;  
 XX 21-JAN-2003 (first entry)  
 XX DNA encoding alpha amylase polypeptide #2.  
 DE Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 XX starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX Unidentified.  
 OS W0200268589-A2.  
 XX 06-SEP-2002.  
 XX 21-FEB-2002; 2002WO-US005068.  
 XX 21-FEB-2001; 2001US-0270495P.  
 XX 21-FEB-2001; 2001US-0270496P.  
 XX 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 XX Callen W, Richardson T, Frey G;  
 WPI; 2003-018656/01.  
 DR P-PSDB; ABU03041.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 1; Fig 16; 301pp; English.  
 PS The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABX08411-ABX08415 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention  
 XX Sequence 1311 BP; 313 A; 378 C; 376 G; 244 T; 0 U; 0 Other;  
 SQ  
 Query Match 87.4%; Score 1146.2; DB 7; Length 1311;  
 Best Local Similarity 92.1%; Pred. No. 8.5e-312;  
 Matches 1208; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
 QY 1 ATGGCCAGTACTCCGAGCTGGAAGGCGGGGTGCTAATGCGAGCGTTCTACTGGGAC 60

Db 1 ATGSCAAAGTACCTGGAGCTGGAAGAGGGCGGGCTCATATGCAAGGCTTCTACTGGGAC 60  
QY 61 GTGCTTTCAGGAGGAATATGTTGGGACACATACGCGAGAGATACCGGAGTGTACGAT 120  
Db 61 GTCCCCATGGGAGGAATCTGTGGGACACGATAGCCAGAGATACCCGACTTGGGCAAGC 120  
QY 121 GCCGGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGSCATGGGCGGCGCTATTTCG 180  
Db 121 GCCGGGATTTCCGGCGATATGATTTCCCGCGGAGCAAGGSCATGGGCGGCGCTATTTCG 180  
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACGAGAGGAGACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACGAGAGGAGACGGTA 240  
QY 241 GAGACGCGCTTTGGCTTCAACGAGGAGCTTCTGAAACATGATAAACAACCGCCACGCTAT 300  
Db 241 GAGACGCGCTTTGGCTTCAACGAGGAGCTTCTGAAACATGATAAACAACCGCCACGCTAT 300  
QY 301 GGCATGAAGGTAATAGCCGATATAGTATCAACCAACCGCGCGCGGTGACCTGAGATGG 360  
Db 301 GGCATGAAGGTAATAGCCGATATAGTATCAACCAACCGCGCGCGGTGACCTGAGATGG 360  
QY 361 AACCCCTTCGTGAAGCACTATACCTGGACCGACTTCTCAAGGTCGCGTGGGTAATATAC 420  
Db 361 AACCCCTTCGTGAAGCACTATACCTGGACCGACTTCTCAAGGTCGCGTGGGTAATATAC 420  
QY 421 ACGGCCAACTACCTCGACTTCCACCCGGAACGAGTCCATCGGCGGATTCGGGAACATTT 480  
Db 421 ACGGCCAACTACCTCGACTTCCACCCGGAACGAGTCCATCGGCGGATTCGGGAACATTT 480  
QY 481 GGAGGCTATCCGACATATGCGACGACGAGCTGGACCGAGTACTGGCTTGGGCGGACG 540  
Db 481 GGAGGCTATCCGACATATGCGACGACGAGCTGGGACCGAGTACTGGCTTGGGCGGACG 540  
QY 541 CAGGAGAGCTACGGGGCATATCTCAGAGAGCATCGGCATCGCTGCGCTTCCGACTAC 600  
Db 541 CAGGAGAGCTACGGGGCATATCTCAGAGAGCATCGGCATCGCTGCGCTTCCGACTAC 600  
QY 601 GTCAGGGCTATGCTCCCTCGGTGCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660  
Db 601 GTCAGGGCTATGCTCCCTCGGTGCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660  
QY 661 GTTGAGAGTACTGGGACCAACAGTCGACGCTGTCTCACTGCGGATCTCGAGCGGT 720  
Db 661 GTTGAGAGTACTGGGACCAACAGTCGACGCTGTCTCACTGCGGATCTCGAGCGGT 720  
QY 721 GCCAAGGCTTTTGACTTCGCTTCTACTCAAGATGGATGAGGCTTTGACAAACAAAAC 780  
Db 721 GCCAAGGCTTTTGACTTCGCTTCTACTCAAGATGGATGAGGCTTTGACAAACAAAAC 780  
QY 781 ATTCAGGCTCGTCTGCTCCCTTCAAGACCGGACGAGCTGTGTCTCCGCGGACCGCTTC 840  
Db 781 ATTCAGGCTCGTCTGCTCCCTTCAAGACCGGACGAGCTGTGTCTCCGCGGACCGCTTC 840  
QY 841 AAGGCGGTAACTTTGTAGCAAAACACGACACCGCATATAATCTGGAACAAAGTATCCAGCC 900  
Db 841 AAGGCGGTAACTTTGTAGCAAAACACGACACCGCATATAATCTGGAACAAAGTATCCAGCC 900  
QY 901 TAGCGGTTTCATCTCCTACGAGGCGGCGGACATATTTCTACCGGAGTACGAGGAG 960  
Db 901 TAGCGGTTTCATCTCCTACGAGGCGGCGGACATATTTCTACCGGAGTACGAGGAG 960  
QY 961 TGGCTCAACAGGATAAGCTCAAGAACTCATCTGGATACATGAGAACCTCGCGCGGAGGA 1020  
Db 961 TGGCTCAACAGGATAAGCTCAAGAACTCATCTGGATACATGAGAACCTCGCGCGGAGGA 1020  
QY 1021 AGCACCGCATAGTCTACTACGATAACGATGAATCTATCTTCGAGGAACGGTACGGG 1080  
Db 1021 AGCACTGACATCGTTTACTACGACAAACGACGAGTGTATTCGTGAGAAACGGTACGGA 1080  
QY 1081 GACAGCGGGGCTTATTAACCTACATCACTTAGCTCGACGAGCGCGGAGTGGTT 1140

Db 1081 AGCAAGCCGGGACTGATACATACATCACTCGCTCAAGCGAAGCCGGAAGGTGGTTC 1140  
QY 1141 TATGTCCCAAGTTCGGCGGCGTGCATCCACGAGTATACCTGTAACCTCGGAGGCTGG 1200  
Db 1141 TACGTTCCGAGTTCCGGGAGCGTGCATCCACGAGTACACCGGACACCTCGCGGCTGG 1200  
QY 1201 GTAGACAAGTACGCTATCAAGCGGCTGGGTATCTCGAAGCTCCAGCTTAAGACCTT 1260  
Db 1201 GTGACAAGTGGGTGGGTCAAGCGGCTGGGTATCTCGAGGCGCTCGGACGACCGG 1260  
QY 1261 GCCAAGCGGCACTATGGCTACTCCGTGAGCTACTCGGGGTGGGCTGA 1311  
Db 1261 GCCAAGCGCTATTACGGCTACTCCGTGAGCTATTGCGGTGTTGGGTGA 1311

RESULT 14

ABX08439  
ID ABX08439 standard; DNA; 1311 BP.  
XX ABX08439;  
XX AC  
XX XX  
XX 21-JAN-2003 (first entry)  
XX XX  
DE DNA encoding alpha amylase polypeptide #29.  
XX Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.

XX Unidentified.  
XX OS  
XX PN WO200268589-A2.  
XX PD 06-SEP-2002.  
XX PF 21-FEB-2002; 2002WO-US005068.  
XX XX  
XX 21-FEB-2001; 2001US-0270495P.  
XX PR 21-FEB-2001; 2001US-0270496P.  
XX PR 14-MAY-2001; 2001US-0291122P.  
XX XX  
XX (DIVE-) DIVERSA CORP.  
XX PA  
XX PI Callen W, Richardson T, Frey G;  
XX XX  
XX WPI; 2003-018656/01.  
XX DR P-PSDB; ABU03068.  
XX XX

PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
PT liquefying starch, for textile desizing, for treating lignocellulosic  
PT fibers, and for producing high-maltose or high-glucose syrup.  
XX XX

Claim 1; Fig 16; 301pp; English.

XX The invention relates to a purified polypeptide with alpha-amylase  
XX activity and the polynucleotide encoding it. The polypeptide is useful  
XX for hydrolysing starch linkages, for catalysing the breakdown of a  
XX starch, for modifying small molecules, for liquefying starch, for washing  
XX an object, for textile desizing, for treating lignocellulosic fibers, for  
XX improving fibre properties, for enzymatic de-linking of recycled paper  
XX pulp, for producing a high-maltose or high-glucose syrup or a mixed  
XX syrup, and for increasing the flow of production fluids from a  
XX subterranean formation by removing a viscous, starch-containing, damaging  
XX fluid formed during production operations and found within the  
XX CC subterranean formation which surrounds a completed well bore. The  
XX polypeptide is also useful for preparing a dough or a baked product  
XX prepared from the dough and in corn-wet milling processes, detergents,  
XX baking processes, beverages, oil fields (fuel ethanol), brewing processes  
XX and starch modification in the paper and pulp industry, for removing  
XX starch containing stains from a material and for reducing staling of

CC bakery products, Sequences ABX08411-ABX08515 represent DNA molecules  
 CC encoding alpha amylase polypeptides of the invention

XX SQ Sequence 1311 BP; 321 A; 361 C; 377 G; 252 T; 0 U; 0 Other;

Query Match 87.4%; Score 1146.2; DB 7; Length 1311;  
 Best Local Similarity 92.1%; Pred. No. 8.5e-312;  
 Matches 1208; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 1 ATGGCCAGTACTCCAGCTGGAAGAGGCGGGGTCAATATCCAGCGTTCCTACTGGAC 60  
 DB 1 ATGGCCAGTACTCCAGCTGGAAGAGGCGGGGTCAATATCCAGCGTTCCTACTGGAC 60  
 QY 61 GTGCTTTACAGGAGGAATATGGTGGGACACAAATACGCGCAAGATACCCGAGTGGTACGAT 120  
 DB 61 GTGCTTTACAGGAGGAATATGGTGGGACACAAATACGCGCAAGATACCCGAGTGGTACGAT 120  
 QY 121 GCGGGAATCTCCGCAATATGGATTTCCCGGCGAGCAAGGGCATGGGCGGCGCTATTTCG 180  
 DB 121 GCGGGAATCTCCGCAATATGGATTTCCCGGCGAGCAAGGGCATGGGCGGCGCTATTTCG 180  
 QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGAGCAAGGTA 240  
 DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGAGCAAGGTA 240  
 QY 241 GAGACGCGCTTTGGCTCCAGGAGGAGCTCGTGAACATGATATAACACCCGCCCTAT 300  
 DB 241 GAAACGAGGTTCGGCTCAAGCAGGAGCTCATAAATGATATAACACCCGCCCTAT 300  
 QY 301 GGCATGAAGGTATAGCCATATAGTATCATCAACCCCGCGCGGTGACTTGGAGTGG 360  
 DB 301 GGCATGAAGGTATAGCCATATAGTATCATCAACCCCGCGCGGTGACTTGGAGTGG 360  
 QY 361 AACCCCTTCGTGAACACTATACCTGGACCGACTTCTCAAGAGTCCGCTCGGTAATAC 420  
 DB 361 AACCCCTTCGTGAACACTATACCTGGACCGACTTCTCAAGAGTCCGCTCGGTAATAC 420  
 QY 421 ACGGCCAACTACTCTGACTTCCACCGGAACGAGTCCATGCGGGGAGATTCGGGAACATTT 480  
 DB 421 ACGGCCAACTACTCTGACTTCCACCGGAACGAGTCCATGCGGGGAGATTCGGGAACATTT 480  
 QY 481 GGAGGCTATCCGACATATGCGACGACGAGCTGGACCACTGCTGGCTCGGCGCAGC 540  
 DB 481 GGAGGCTATCCGACATATGCGACGACGAGCTGGACCACTGCTGGCTCGGCGCAGC 540  
 QY 541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
 DB 541 CAGGAGGCTACGCGGCATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
 QY 601 GTCAAGGCTATGCTCCCTGGTTCGTCAGGACTGGCTGAACCTGGTGGGAGCTGGCG 660  
 DB 601 GTCAAGGCTATGCTCCCTGGTTCGTCAGGACTGGCTGAACCTGGTGGGAGCTGGCG 660  
 QY 661 GTTGAGAGTACTGGGACACCAACGCTCGACGCTGTTCTCAACTGGGCATCTCGAGCGGT 720  
 DB 661 GTTGAGAGTACTGGGACACCAACGCTCGACGCTGTTCTCAACTGGGCATCTCGAGCGGT 720  
 QY 721 GCCAAGCTTTTGACTTCGCTCCTACTACAGATGATGAGGCTTTGACACCAAAAC 780  
 DB 721 GCCAAGCTTTTGACTTCGCTCCTACTACAGATGATGAGGCTTTGATACCAACAC 780  
 QY 781 ATTCCAGCGCTGCTCTGCTCCCTTCAGAACCGCCAGACTGTTGTCTCCCGCAGCCGTTTC 840  
 DB 781 ATTCCAGCGCTGCTGCTCCCTTCAGAACCGCCAGACTGTTGTCTCCCGCAGCCGTTTC 840  
 QY 841 AAGCGCGTAACTTTGTAGCAACCAACGACCGATATATCTTGAAACAGATATCCAGCC 900  
 DB 841 AAGCGCTGTGACGCTTTGTAGCAACCAACGATACCGATATAATTTGGAACAGATATCCAGCC 900  
 QY 901 TAGCGGTTCTATCTCCTACCTACGAGGCGCCAGCAATATTTCTACCGGACTACGAGGAG 960  
 DB 901 TAGCGGTTCTATCTCCTACCTACGAGGCGCCAGCAATATTTCTACCGGACTACGAGGAG 960

QY 961 TGGCTCAACAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGAGGA 1020  
 DB 961 TGGCTCAACAAGGACAGGCTCAAGAACCTCATCTGGATACACGACCACTCGCCGAGGA 1020  
 QY 1021 AGCACCGACATAGTCTACTACGATAACGATGAACACTCATCTTGTACGGAACGGCTACGGG 1080  
 DB 1021 AGCACTGACATCTGTTTACTACGACAACGAGAGCTGATATTCGTGAGAAACGGCTACGGA 1080  
 QY 1081 GACAAGCCGGGCTTAACTACATCAACCTAGGCTCGAGCAAGCCGGAAGTGGT 1140  
 DB 1081 AGCAAGCCGGGCTGATACATCAACCTAGGCTCGAGCAAGCCGGAAGTGGT 1140  
 QY 1141 TATGTGCGGAAGTTCGCGGCGGTGCATCCACGAGTACTGGTAACTCGGAGGCTGG 1200  
 DB 1141 TATGTGCGGAAGTTCGCGGCGGTGCATCCACGAGTACTGGTAACTCGGAGGCTGG 1200  
 QY 1201 GTAGACAGTACCTCTACTCAAGGGCTGGGTCTATCTGMAAGCTCCAGCTTACGACCT 1260  
 DB 1201 GTAGACAGTACCTCTACTCAAGGGCTGGGTCTATCTGMAAGCTCCAGCTTACGACCT 1260  
 QY 1261 GCCAAACGGGCGAGTATGGTACTCCGTGTGGAGCTACTGCGGGGTGGGCTCA 1311  
 DB 1261 GCCAAGCGGCGAGTATGGTACTCCGTGTGGAGCTATTCGGGTGTTGGGTGA 1311

## RESULT 15

ABX08435  
 ID ABX08435 standard; DNA; 1311 BP.

XX AC ABX08435;

XX AC

DT 21-JAN-2003 (first entry)

XX DT

DE DNA encoding alpha amylase polypeptide #25.

XX DE

XX KW Alpha amylase; gene; ds; starch linkage hydrolysis; starch liquefaction;  
 starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 high glucose syrup; corn-wet milling process; detergent; baking process;  
 beverage; oil field; fuel ethanol; brewing process; staling;  
 starch modification.

XX KW

OS Unidentified.

XX OS

XX FN WO200268589-A2.

XX FN

PD 06-SEP-2002.

XX PD

PF 21-FEB-2002; 2002WO-US005068.

XX PF

PR 21-FEB-2001; 2001US-0270495P.

XX PR

PR 21-FEB-2001; 2001US-0270496P.

XX PR

PR 14-MAY-2001; 2001US-0291122P.

XX PR

PA (DIVE-) DIVERSA CORP.

XX PA

PI Callen W, Richardson T, Frey G;

XX PI

XX WPI; 2003-018656/01.

DR P-PSDB; ABU03064.

XX DR

XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 liquefying starch, for textile desizing, for treating lignocellulosic  
 fibers, and for producing high-maltose or high-glucose syrup.

PT PT

PT PT

XX PT

XX Claim 1; Fig 16; 301pp; English.

FS FS

XX FS

CC The invention relates to a purified polypeptide with alpha-amylase  
 activity and the polynucleotide encoding it. The polypeptide is useful  
 for hydrolysing starch linkages, for catalysing the breakdown of a  
 starch, for modifying small molecules, for liquefying starch, for washing  
 an object, for textile desizing, for treating lignocellulosic fibers, for  
 improving fibre properties, for enzymatic de-inking of recycled paper

CC CC



CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging  
CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing scaling of  
CC bakery products. Sequences ABX08411-ABX08515 represent DNA molecules  
XX encoding alpha amylase polypeptides of the invention

Sequence 1311 BP; 317 A; 371 C; 378 G; 245 T; 0 U; 0 Other;

Query Match 87.3%; Score 1144.6; DB 7; Length 1311;  
Best Local Similarity 92.1%; Pred. No. 2.4e-311;  
Matches 1207; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 1 ATGGCAAAGTATCTCGAGCTCGAAGAGGCGGGGTCAATATGACAGGCGTCTACTGGGAC 60  
DB 1 ATGGCAAAGTATCTCGAGCTCGAAGAGGCGGGGTCAATATGACAGGCGTCTACTGGGAC 60

QY 61 GTGGCTTCAGAGGAATATGTGGGACACAATAGCGGAGAGATACCGGAGTGTAGCAT 120  
DB 61 GTGGCTTCAGAGGAATATGTGGGACACAATAGCGGAGAGATACCGGAGTGTAGCAT 120

QY 121 GCCGAAATCTCGCAATATGATTTCCCGCGGAGCAGGCGGCGGCGCTATTTCG 180  
DB 121 GCCGAAATCTCGCAATATGATTTCCCGCGGAGCAGGCGGCGGCGCTATTTCG 180

QY 181 ATGGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGTA 240  
DB 181 ATGGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGTA 240

QY 241 GAGAGCGCTTTGGCTCAAGCAGGAGCTCGTGACATGATAAACCAGCCGACGCTAT 300  
DB 241 GAGAGCGCTTTGGCTCAAGCAGGAGCTCGTGACATGATAAACCAGCCGACGCTAT 300

QY 301 GGCAATGAAGTAATAGCGGATATAGTCAATCAACACCGCGCGGCTGACCTGGAGTGG 360  
DB 301 GGCAATGAAGTAATAGCGGATATAGTCAATCAACACCGCGCGGCTGACCTGGAGTGG 360

QY 361 AACCCCTTCGTGAACGATATACCTGGACCGACTTCTCAAAAGTTCGGGTAAATAC 420  
DB 361 AACCCCTTCGTGAACGATATACCTGGACCGACTTCTCAAAAGTTCGGGTAAATAC 420

QY 421 ACGGCAACTACCTCGACTTCCACCGAAGCAGCTCCATGCGGCGATTCCGGAACATTT 480  
DB 421 ACGGCAACTACCTCGACTTCCACCGAAGCAGCTCCATGCGGCGATTCCGGAACATTT 480

QY 481 GGAGCTATCCCGACATATGCCACGACAGAGCTGGGACCAAGTACTGGCTCTGGGCCAGC 540  
DB 481 GGAGCTATCCCGACATATGCCACGACAGAGCTGGGACCAAGTACTGGCTCTGGGCCAGC 540

QY 541 CAGGAGAGTACCGGCGATATCTCAGGAGCATCGGCATCGATCGCTGGCGCTTCGACTAC 600  
DB 541 GATGAGAGCTACGCGCCTTACCTAAGGAGCATCGGCGTTGATGCTGGCGCTTTGACTAC 600

QY 601 GTCAAGGGCTATGCTCCCTCGGTCTGCAAGGACTCGCTGAACTGGTGGGAGGCTGGGCG 660  
DB 601 GTGAGGGCTACGAGCGTGGGTCTGCAAGGACTGGCTCACTGGTGGGCGGCTGGGCC 660

QY 661 GTTGAGAGTACTGGGACACCAACGTGACGCTGTTCTCAACTGGGCGATACTCGAGCGGT 720  
DB 661 GTTGAGAGTACTGGGACACCAACGTGACGCTGTTCTCAACTGGGCGCTACTCGAGCGGC 720

QY 721 GCCAAGGCTTTGACTTCGCGCTCTACTACAGATGGATGAGGCTTTGACAAACAAAC 780  
DB 721 GCCAAGGCTTTGACTTCGCGCTCTACTACAGATGGATGAGGCTTTGACAAACAAAC 780

QY 781 ATTCAGCGCTCGTCTCTGCCCTTCAGAACCGGCCAGACTGTGTCTCCCGCAGCCGTTTC 840

DB 781 ATTCAGCGCTCGTCTCTGCCCTTCAGAACCGGCCAGACTGTGTCTCCCGCAGCCGTTTC 840  
QY 841 AAGGCGGTAACTTTGTAGCAAAACACGACACGGATATAATCTGGAAACAAGTATCCAGCC 900  
DB 841 AAGGCGGTAACTTTGTAGCAAAACACGACACGGATATAATCTGGAAACAAGTATCCAGCC 900  
QY 901 TAGCGCTTCATCTCTACCTACGAGGCGCAGCGACAATATTCTACCGCGACTACGAGGAG 960  
DB 901 TAGCGCTTCATCTCTACCTACGAGGCGCAGCGACAATATTCTACCGCGACTACGAGGAG 960  
QY 961 TGCTCTCAACAAGGATTAAGCTCAAGAACCTCATCTGGATACATGAGAACCTTCGCGCGAGGA 1020  
DB 961 TGCTCTCAACAAGGATTAAGCTCAAGAACCTCATCTGGATACATGAGAACCTTCGCGCGAGGA 1020

QY 1021 AGCACCGACATAGTCTACTACGATACGATGAACATCTTCTGTCAGGAACGCTACGGG 1080  
DB 1021 AGCACCGACATAGTCTACTACGATACGATGAACATCTTCTGTCAGGAACGCTACGGG 1080

QY 1081 GACAAGCCGCGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGGCCCGGAAAGGTGGTT 1140  
DB 1081 GACAAGCCGCGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGGCCCGGAAAGGTGGTT 1140

QY 1141 TATGTCCGAAAGTTCGCGGCGGCTGATCCAGGAGTATCTGTTAACTTCGAGGCTGG 1200  
DB 1141 TATGTCCGAAAGTTCGCGGCGGCTGATCCAGGAGTATCTGTTAACTTCGAGGCTGG 1200

QY 1201 GTAGACAAGTACGTCTPACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
DB 1201 GTAGACAAGTACGTCTPACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

QY 1261 GCCAAGCGGCGAGTATGGCTACTCCGTTGTGGAGCTACTCGGGGTGGGCTGA 1311  
DB 1261 GCCAAGCGGCGAGTATGGCTACTCCGTTGTGGAGCTATTGGGCTGTTGGCTGA 1311

Search completed: June 29, 2004, 09:14:42  
Job time : 578 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: June 29, 2004, 09:04:55 ; Search time 3765 Seconds  
(without alignments)  
10398.225 Million cell updates/sec

Title: US-10-081-739A-1  
Perfect score: 1311  
Sequence: 1 atggccagactccgagct.....gctactgcgggtgggctga 1311

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107.6	8.2	540	9 AUI92670	AUI92670 AUI92670
2	82.8	6.3	428	13 BU639413	BU639413 mgcw013xn
3	76.8	5.9	1628	28 B2424843	B2424843 100020232
4	72.8	5.6	394	13 BU640392	BU640392 mgcw019x0

5	71.2	5.4	768	28	A0159619	A0159619 mgxb0001K
6	63.2	4.8	660	14	CD900121	CD900121 GL74.1141
7	62.6	4.8	681	29	CNS0100L	AL153781 Anopheles
8	62.4	4.8	858	29	CG302842	CG302842 OG2AE22TV
9	60.8	4.6	691	29	CG213284	CG213284 OG2AEW8TH
10	58	4.4	519	14	CB884205	CB884205 HVI0M08r
11	56.6	4.3	422	9	AL820356	AL820356 AL820356
12	54	4.1	593	29	CG289556	CG289556 OG5AR78TC
13	54	4.1	602	29	CG345046	CG345046 OGVDL11TH
14	54	4.1	608	29	CG222587	CG222587 OGWF50TH
15	54	4.1	695	28	CC344571	CC344571 OG0BE54TV
16	53.6	4.1	559	14	CF322783	CF322783 HDN--02-B
17	53	4.0	405	14	CF323742	CF323742 HDN--04-K
18	52.8	4.0	562	12	BJ477953	BJ477953 BJ477953
19	52.6	4.0	558	14	CF323939	CF323939 HDN--05-C
20	52.4	4.0	474	10	AW923088	AW923088 DGL_50_H1
21	52	4.0	743	28	AQ156694	AQ156694 mgx0001I
22	51.6	3.9	841	12	B1949299	B1949299 HVSME1001
23	51.2	3.9	613	29	CC639582	CC639582 OGKAK19TH
24	51.2	3.9	982	28	CC424389	CC424389 PUHNS3TB
25	48.8	3.7	681	29	CNS02E0D	AL193990 Tetraodon
26	46.6	3.6	424	10	BF177257	BF177257 EM1_1_B05
27	46.6	3.6	451	12	BG556788	BG556788 EM1_38_F0
28	46.6	3.6	532	12	BG557843	BG557843 EM1_55_A1
29	46.6	3.6	541	10	BF177143	BF177143 EM1_2_C09
30	46.6	3.6	542	10	BE361447	BE361447 DGL_72_D0
31	46.6	3.6	545	12	BG355757	BG355757 EM1_19_B1
32	46.6	3.6	615	10	BE360090	BE360090 DGL_61_C1
33	46.6	3.6	936	29	CG038303	CG038303 PUFNQ30TB
34	46.2	3.5	500	9	AA898808	AA898808 NCP4C8T3
35	45.8	3.5	474	10	BE249153	BE249153 NCP38H08D
36	45.8	3.5	584	14	CF554700	CF554700 1115001H0
37	45.8	3.5	869	14	CK159167	CK159167 FGAS04056
38	45.6	3.5	740	13	BQ801516	BQ801516 WHE2815_C
39	44.8	3.4	525	14	CA700958	CA700958 wkm1c.DK0
40	44.8	3.4	527	10	BE293754	BE293754 WHE2154_F
41	44.8	3.4	834	29	CG133420	CG133420 PU1AL15TD
42	44.8	3.4	1201	13	BX356664	BX356664 BX356664
43	44.6	3.4	404	10	AW923211	AW923211 DGL_50_G0
44	44.6	3.4	642	14	CF306714	CF306714 HDAL--04-
45	44.6	3.4	1201	13	BX381961	BX381961 BX381961

ALIGNMENTS

RESULT 1  
AUI92670  
LOCUS AUI92670 540 bp mRNA linear EST 14-OCT-2003  
DEFINITION Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis  
cdna clone PFL034G01\_r 5', mRNA sequence.  
ACCESSION AUI92670  
VERSION AUI92670.1 GI:31931544  
KEYWORDS EST.  
SOURCE Porphyra yezoensis  
ORGANISM Porphyra yezoensis  
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
REFERENCE 1 (bases 1 to 540)  
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.  
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS  
JOURNAL J. Phycol. 39 (5), 923-930 (2003)  
COMMENT Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/-/Location/Qualifiers  
1. 540  
/organism="Porphyra yezoensis"

```

/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL03401.r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"

```

## ORIGIN

```

Query Match      8.2%; Score 107.6; DB 9; Length 540;
Best Local Similarity 58.4%; Pred. No. 1.1e-13;
Matches 188; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 30 CGGGGTCAATAGTCAGCGCTTCTACTGGACCTGCTTCAGGAGAAATGTTGGGACAC 89
DB 217 CGGGGTCTCTCCAGTCTTCTACTGGACCTTCCCGCGGATGGAGCTGTGGAAGAA 276

QY 90 AATACGGCAGAGATACCGGAGTGTGTACGATCCGGAATCTCCGCAATATGATTCCCCC 149
DB 277 CTTGGCGGTGAACGGGTGGACCTCGCGCGCGGGCGTCACTGCGTCTGGCTGCGCC 336

QY 150 GCGGACGAGGATGCGGGCGGCGCTTCTCGATGGCTACGACCCCTAGGACTTCTTTGA 209
DB 337 GCGGTACAAAGGGCGAGCGCGGGCGCGCGGCGTGGCTATGGCTGTACGACACGATATGA 396

QY 210 CTTGGGTGATGACGACCAAGGAGGAGCGGTAGAGACGCGCTTTGGCTCCAGCAGGAGCT 269
DB 397 CTTGGGGAGTTTGACCAAAAGGGCACCGTACGGACCAAGTACGACCAAGGAGCAGCT 456

QY 270 CGTGAACATGATAAACACCGCGCCAGCCCTATGGCATGAGGTAAATAGCGGATATAGTCAT 329
DB 457 CTTGTGCGCGCGCTCGCGCGCGCATGACGCGGCGCATCCAGGTGTACGCGGACATTTGTCT 516

QY 330 CAACACCGCGCGCGGGGTGAC 351
DB 517 CAACACCGGGTGGCGGAC 538

```

```

RESULT 2
BU639413
LOCUS      428 bp mRNA linear EST 06-MAY-2003
DEFINITION mgcw013xN24f.b RCW Lambda Zap Express Library Magnaporthe grisea
ACCESSION  BU639413
VERSION     BU639413.2 GI:30396874
KEYWORDS   EST.
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS    Sordariomycetes Incertae sedis; Magnaportheaceae; Magnaporthe.
           1. (bases 1 to 428)
           Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P., Lu,G.,
           Bhatterai,K. and Dean,R.A.
TITLE      Expressed sequence tags from the rice blast fungus, Magnaporthe
           grisea
JOURNAL    Unpublished (2002)
COMMENT    On Sep 30, 2002 this sequence version replaced gi:23351739.
           Contact: Ebbole D
           Department of Plant Pathology & Microbiology
           Texas A&M University
           Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
           Tel: 979 845 4831
           Fax: 979 845 6483
           Email: d-ebbole@tamu.edu
           Chromatogram file of this sequence is available, see contact
           person:Best nr hit (April. 22, 2003) gb|EAA33974.1| hypothetical
           protein [Neurospora crassa] 178 1e-44
PCR PRIMERS
FORWARD: T3 primer
BACKWARD: T7 primer
Place: mgcw013 row: N column: 24
Seq primer: T3.

```

## FEATURES

```

location/Qualifiers
1..428

```

```

/organism="Magnaporthe grisea"
/mol_type="mRNA"
/strain="Cp987"
/db_xref="taxon:149305"
/clone="mgcw013xN24"
/sex="Mat-1 hermaphrodite"
/tissue_type="Mycelium"
/dev_stage="Day 5 post-inoculation"
/clone_lib="RCW Lambda Zap Express Library"
/notes="Vector: pBluescript excised from Lambda Zap
Express; Site 1: EcoRI; Site 2: XhoI; Day 5
post-inoculation MRNAs prepared from Magnaporthe grisea
grown at 23C in the dark with constant gyratory shaking
100 rpm in Vogel's minimal medium containing 0.5% isolated
rice cell walls as the sole carbon source. Library
provided by Sheng-Cheng Wu. Sequences were processed by
one of two methods. Where a full-length alignment to the
M. grisea genome sequence was available, the EST sequence
was trimmed according to the alignment, otherwise sequence
quality was assessed using phredPhrap version 991019 and
trimmed according to phd files (0.05) and for vector
seqs."

```

## ORIGIN

```

Query Match      6.3%; Score 82.8; DB 13; Length 428;
Best Local Similarity 55.5%; Pred. No. 4.2e-08;
Matches 181; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 37 ATATGCGAGCGTTCTACTGGGACGTGCTTCTAGGAGGAATATGTGGGACACAAATACGG 96
DB 63 ATATGCGAGCGTTCTAGTGGTATGTCCTCCAGCAGATCAGAAACACTGGTCTGAGTTGAG 122

QY 97 CAGAAGATACCGGAGTGGTACGATCCGGAATCTCCGCAATATGGATTCCCGCGGGAGC 156
DB 123 AAGGATATCCCGAGCTCAATCATGGGTATGCACAATATCTGGTTCCACAGGATGC 182

QY 157 AAGGCGATGGCGCGCGCTTATTCGATGGGTACGACCCCTACGACTTTTTCACCTCGT 216
DB 183 AAGGCTCGTCCAAAGACCGGC---AATGGTACGACATCTATGACCTCTATGACCTGGGC 239

QY 217 GAGTACGACGAGGAGGACGCTAGAGACGCTTTGGCTCCAGCAGGAGCTCTGTAAC 276
DB 240 GAGTTCCACCAAAAGGAAAGCGTGGCCACCAATATGGGAAACCAAGGAGGAGCTCTGTAAG 299

QY 277 ATGATAAACACCCCGACGCTATGTCATGAAGGTAAATAGCCGATATAGTCTCAACACAC 336
DB 300 CTTTGTCTACGGCCAAAGCCGTTGGCATCTACTGGGACCGGTTCTGAACACAC 359

QY 337 CGCGCGCGGCTGACCTGGAGTGAA 362
DB 360 AGGTTTTCAGCCGACCAATGAAGGAA 385

```

## RESULT 3

```

BU624843/c
LOCUS      1628 bp DNA linear GSS 13-DEC-2002
DEFINITION 100020232-5732 Aspergillus terreus random genomic DNA clone library
Aspergillus terreus genomic, genomic survey sequence.
ACCESSION  BU624843
VERSION     BU624843.1 GI:26666298
KEYWORDS    GSS.
SOURCE      Aspergillus terreus
ORGANISM    Aspergillus terreus
REFERENCE  1 (bases 1 to 1628)
AUTHORS     Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
           Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
           1 (bases 1 to 1628)
           Askenazi,M., Driggers,E.M., Holtzman,D.A., Norman,T.C., Iverson,S.,
           Zimmer,D.P., Boers,M.E., Blomquist,P.R., Martinez,E.J.,
           Monreal,A.W., Feibelman,T.P., Mayorga,M.E., Maxon,M.E., Sykes,K.,
           Tobin,J., Cordero,E., Salama,S.R., Trueheart,J., Royer,J.C. and
           Madden,K.T.
TITLE      Integrating transcriptional and metabolite profiles to direct the
           engineering of lovastatin-producing strains

```

JOURNAL COMMENT  
Unpublished (2002)  
Contact: Zimmer DP  
Microbia, Inc.  
One Kendall Square Building 1400 W, Cambridge, MA 02139, USA  
Tel: 617-621-8322  
Fax: 617-  
Email: dzimmer@microbia.com  
Class: plasmid ends.  
Location/Qualifiers  
1. 1628  
/organism="Aspergillus terreus"  
/mol\_type="genomic DNA"  
/strain="ATCC 20542 (A. terreus Thom, anamorph)"  
/db\_xref="taxon:33178"  
/lab\_host="Escherichia coli"  
/clone\_lib="Aspergillus terreus random genomic DNA clone library"  
/note="Vector: pZEROTM-2; Site 1: Sau3A; Site 2: BamHI; Sau3A genomic fragments ligated into BamHI digested pZEROTM-2"

FEATURES  
source  
1. 1628  
Location/Qualifiers  
/organism="Magnetorhys grisea"  
/mol\_type="rRNA"  
/strain="CP987"  
/db\_xref="taxon:148305"  
/clone="mgcw019x009"  
/sex="Mat1-1 hermaphrodite"  
/tissue\_type="Mycelium"  
/dev\_stage="Day 5 post-inoculation"  
/clone\_lib="RCW Lambda Zap Express Library"  
/note="Vector: pBluescript excised from Lambda Zap Express; Site 1: EcoRI; Site 2: XhoI; Day 5 post-inoculation mRNAs prepared from Magnetorhys grisea grown at 23C in the dark with constant gyratory shaking 100 rpm in Vogel's minimal medium containing 0.5% isolated rice cell walls as the sole carbon source. Library provided by Sheng-Cheng Wu. Sequences were processed by one of two methods. Where a full-length alignment to the M. grisea genome sequence was available, the EST sequence was trimmed according to the alignment, otherwise sequence quality was assessed using phredPhrap version 991019 and trimmed according to phd files (0.05) and for vector seqs."

ORIGIN  
Query Match 5.9%; Score 76.8; DB 28; Length 1628;  
Best Local Similarity 54.7%; Pred. No. 1.7e-06;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 33 GGTATATGACGGCTTCTACTGGACGTCCTTCAGGAGATATGGTGGACACAAT 92  
Db 317 GCTAATCTTCCAGGGGTTGATGGCAGCTCCCGGACAGCGGAGGCACTGGCGGGCT 258  
QY 93 ACGCAGAGATACCGGAGTGGTACGATCGCGGAATCTCCGCAATATGGATTCCCGCGC 152  
Db 257 CGCGGGCATTCGCGGATTAAGGCCATCGGAATGCACTACATATGGCTTCCCGCGG 198  
QY 153 GAGCAAGGCGATGGCGCGGCTTATCGATGGGTACGACCCCTACGACTTCTTTGACCT 212  
Db 197 ATGCAAGGGATGGACC---CATCGGGGACCGGTATGATATCTATATGACTT 141  
QY 213 CGGTGAGTACACAGACAGGAGGACGTTAGACGCGCTTGGCTCCAGAGGAGCTCGT 272  
Db 140 GGGGAGTTCCATCAGAAAGCTTATCAGCAGAAATGGGGGACACGGGAGGAGCTGCA 81  
QY 273 GAAATATGATAAACCGCCCGCTATGGCATGAAGGTATATGCGCATATAGTCATCAA 332  
Db 80 CGAGCTAGTTAACGACGCTCAGCCCTCGGCTGGGGTCTATTGGGACCGGTTCTCAA 21  
QY 333 CCACGGCGCGGGTGACC 352  
Db 20 CCACAGGCTGGTGGGATC 1

RESULT 4  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Magnetorhys grisea (anamorph: Pyricularia grisea)  
Magnetorhys grisea  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.  
1 (bases 1 to 394)  
Ebbole D.J., Yuan, J., Thomas, T.L., Bobrowicz, P., Lu, G.,  
Bhatterai, K. and Dean, R.A.  
Expressed sequence tags from the rice blast fungus, Magnetorhys grisea  
Unpublished (2002)  
On Sep 30, 2002 this sequence version replaced gi:23352718.  
Contact: Ebbole DJ  
Department of Plant Pathology & Microbiology  
Texas A&M University

Peterson Bldg, MS2132, College Station, TX 77843-2132, USA  
Tel: 979 845 4831  
Fax: 979 845 6483  
Email: d-ebbole@tamu.edu  
Chromatogram file of this sequence is available, see contact  
person: Best nr hit (April, 22, 2003) gb|EAA33974.1| hypothetical  
protein [Neurospora crassa] 155 9e-38  
PCR Primers  
FORWARD: T3 primer  
BACKWARD: T7 primer  
Plate: mgcw019 row: O column: 09  
Seq primer: T3  
Location/Qualifiers  
1. 394  
/organism="Magnetorhys grisea"  
/mol\_type="rRNA"  
/strain="CP987"  
/db\_xref="taxon:148305"  
/clone="mgcw019x009"  
/sex="Mat1-1 hermaphrodite"  
/tissue\_type="Mycelium"  
/dev\_stage="Day 5 post-inoculation"  
/clone\_lib="RCW Lambda Zap Express Library"  
/note="Vector: pBluescript excised from Lambda Zap Express; Site 1: EcoRI; Site 2: XhoI; Day 5 post-inoculation mRNAs prepared from Magnetorhys grisea grown at 23C in the dark with constant gyratory shaking 100 rpm in Vogel's minimal medium containing 0.5% isolated rice cell walls as the sole carbon source. Library provided by Sheng-Cheng Wu. Sequences were processed by one of two methods. Where a full-length alignment to the M. grisea genome sequence was available, the EST sequence was trimmed according to the alignment, otherwise sequence quality was assessed using phredPhrap version 991019 and trimmed according to phd files (0.05) and for vector seqs."

FEATURES  
source  
1. 394  
Location/Qualifiers  
/organism="Magnetorhys grisea"  
/mol\_type="rRNA"  
/strain="CP987"  
/db\_xref="taxon:148305"  
/clone="mgcw019x009"  
/sex="Mat1-1 hermaphrodite"  
/tissue\_type="Mycelium"  
/dev\_stage="Day 5 post-inoculation"  
/clone\_lib="RCW Lambda Zap Express Library"  
/note="Vector: pBluescript excised from Lambda Zap Express; Site 1: EcoRI; Site 2: XhoI; Day 5 post-inoculation mRNAs prepared from Magnetorhys grisea grown at 23C in the dark with constant gyratory shaking 100 rpm in Vogel's minimal medium containing 0.5% isolated rice cell walls as the sole carbon source. Library provided by Sheng-Cheng Wu. Sequences were processed by one of two methods. Where a full-length alignment to the M. grisea genome sequence was available, the EST sequence was trimmed according to the alignment, otherwise sequence quality was assessed using phredPhrap version 991019 and trimmed according to phd files (0.05) and for vector seqs."

ORIGIN  
Query Match 5.6%; Score 72.8; DB 13; Length 394;  
Best Local Similarity 56.9%; Pred. No. 7.4e-06;  
Matches 153; Conservative 0; Mismatches 113; Indels 3; Gaps 1;  
QY 37 ATAATCAGCGCTTCTACTGGGACGTCCTTCAGGAGGAATATGGTGGACACAATACGG 96  
Db 91 ATGATCAGGGCTTTGAGTGGTATGTCCAGCAGATCAGAAACACTGGGTTCAGGTTGAG 150  
QY 97 CAGAAATACCGAGTGGTACGATGCCGAATCTCGCAATATGGATTCCTCCCGCGGACG 156  
Db 151 AAGGAGATCCCCAGCTCANATCATGGGTATCGACAATATCTGGGTTCACAGGATGC 210  
QY 157 AAGGGCATGGCGGCGCTTATTCGATGGGCTACGACCCCTACGACTTCTTTGACCTCGGT 216  
Db 211 AAGGGCTCTCCAAGACCGGC---AATGGGTACGACATCTATGACCTCCATGACCTGGGC 267  
QY 217 GAGTACGACAGAGGAGACGGTAGAGCGCGTTTGGCTTCAAGCAGGAGCTCGTGAAC 276  
Db 268 GAGTTCGACCAAAAGGAGGAGCGTGGCCACCAATGTTGAACCAAGAGGAGCTCGTCAAG 327  
QY 277 ATGATAAACACCGCCACCGCTTATGGCAT 305  
Db 328 CTTTGTCTTACGGTCAAGGCCAAGCGCT 356

RESULT 5  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Magnetorhys grisea (anamorph: Pyricularia grisea)  
Magnetorhys grisea  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.  
1 (bases 1 to 394)  
Ebbole D.J., Yuan, J., Thomas, T.L., Bobrowicz, P., Lu, G.,  
Bhatterai, K. and Dean, R.A.  
Expressed sequence tags from the rice blast fungus, Magnetorhys grisea  
Unpublished (2002)  
On Sep 30, 2002 this sequence version replaced gi:23352718.  
Contact: Ebbole DJ  
Department of Plant Pathology & Microbiology  
Texas A&M University

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariomycetes, incertae sedis; Magnaporthaceae; Magnaporthe.  
 1 (bases 1 to 768)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
 Genome  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: GGAAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence stop: 400.  
 Location/Qualifiers  
 1..768  
 /organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgx50001K09r"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice Blast BAC Library"  
 /notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;  
 Rice blast is one of the most devastating fungal diseases  
 of rice world wide. It is a filamentous ascomycete with  
 a haploid genome (n=7) of approximately 40 Mbp. Rice  
 blast is an important model fungal pathogen for studying  
 numerous aspects of the fungal-host interaction. In  
 order to facilitate genome wide analysis, a BAC library  
 containing 9216 clones with an average insert size of 130  
 kbp was constructed. This library represents greater  
 than 25X genome coverage. High density colony filters  
 are available upon request."

# FEATURES

source

1..768

/organism="Magnaporthe grisea"

/mol\_type="genomic DNA"

/strain="70-15"

/db\_xref="taxon:148305"

/clone="mgx50001K09r"

/tissue\_type="Protoplasts"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice Blast BAC Library"

/notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;  
 Rice blast is one of the most devastating fungal diseases  
 of rice world wide. It is a filamentous ascomycete with  
 a haploid genome (n=7) of approximately 40 Mbp. Rice  
 blast is an important model fungal pathogen for studying  
 numerous aspects of the fungal-host interaction. In  
 order to facilitate genome wide analysis, a BAC library  
 containing 9216 clones with an average insert size of 130  
 kbp was constructed. This library represents greater  
 than 25X genome coverage. High density colony filters  
 are available upon request."

# ORIGIN

Query Match 5.4%; Score 71.2; DB 28; Length 768;  
 Best Local Similarity 58.5%; Pred. No. 2.3e-05;  
 Matches 124; Conservative 0; Mismatches 88; Indels 0; Gaps 0;  
 QY 139 TGGATTTCCCGCGGAGGCGCATGGCGCGGCTATTTCGATGGGCTACGACCCCTAC 198  
 Db 2 TGGATCCACCGGCTGTCAAGGCGGCGGTGGTCTGGAGCAACGGCTACGACGTCTAC 61  
 QY 199 GACTCTTTTGACTCGGTGAGTACGACGAGAGGAAACGGTAGACGCGCTTTGGCTCC 258  
 Db 62 GATCTGTATGATCTGGGCGAGTTTACCAAGAAAGGCTCCAAAGGCGCAAGTGGGGGCCCC 121  
 QY 259 AAGCAGGAGCTGTGAACATGATAACACCGCCGCTATGCGATGAAGTAAATAGCC 318  
 Db 122 AGGACGACCTGGACGAGCTCGTCGCGCGCGCGGGGAGCGCGGATTCGAGATCTGTATT 181  
 QY 319 GATATAGTATCATCAACACCGCGCGCGGTGA 350  
 Db 182 GATCGCGTGTGAACACCAAGGCGCGGCGGA 213

# RESULT 6

CD900121  
 LOCUS  
 DEFINITION  
 G174.114118F01.0824 G174 Triticum aestivum cDNA clone G174114118,  
 mRNA sequence.  
 ACCESSION  
 CD900121  
 VERSION  
 CD900121.1 GI:32674449  
 KEYWORDS  
 EST.  
 SOURCE  
 Triticum aestivum (bread wheat)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 660)  
 Genoplatte.  
 Genoplatte, a major partnership french program in plant genomics  
 Unpublished (2003)  
 Contact: Genoplatte  
 Genoplatte  
 93, rue Henri Rochefort 91025 EVRY CEDEX France  
 Tel: 33 1 69 47 54 00  
 Fax: 33 1 69 47 54 10  
 This sequence has been generated in the framework of the french  
 plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>  
 and <http://genoplatte-info.infobiogen.fr>).

# FEATURES

source

1..660

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="recital"

/db\_xref="taxon:4565"

/clone="G174114118"

/tissue\_type="grain (174 degrees per day after  
 pollination)"

/clone\_lib="G174"

# ORIGIN

Query Match 4.8%; Score 63.2; DB 14; Length 660;  
 Best Local Similarity 46.7%; Pred. No. 0.0014;  
 Matches 200; Conservative 0; Mismatches 228; Indels 0; Gaps 0;  
 QY 217 GAGTAGACACAGAGGAGGAGGAGAGACGGCTTTGGCTCCAGGAGGAGCTCGTGAAC 276  
 Db 4 GCGTCCGGGCTCTACGACATCGACGGCTCCAGTACGCAACGCGGCGGAGCTCAAGTCG 63  
 QY 277 ATGATTAACACCGCCCGGCTATGGCATGAAGGTAATAGCCGATATAGTCATCAACAC 336  
 Db 64 CTATCGCGCGCTCCAGGCAAGGGGTCAGGCGATCGCGACATCTCATCAACAC 123  
 QY 337 CGCGCGCGGCTGACCTGGAGTGGAAACCGCTTCGTGAAGACTATACCTGGACCGGCTTC 396  
 Db 124 CGCTGCGCGGACTACAAGGACAGCGCGGCTCTACTGCGATCTTCGAGGCGCGGACCTCC 183  
 QY 397 TCAAGTTCGCTCGGGTAAATACAGGCAACCTAGCTTCCACCGGAGGCTC 456  
 Db 184 GACGCGCGCTCGACTGGGGCCCCACATGATTCGCGGAGACACACCTACTCCGAC 243  
 QY 457 CATGCGGGCGATTCGCGAACATTTGGAGGCTATCCGACATATGCCACGACAGAGCTGG 516  
 Db 244 GGCACCGCAACCTCGACACCGCGCGGCTTCGCGCGCGCGCGGACATCGACACCTC 303  
 QY 517 GACAGTATCGGCTCTGGGCGGACCGAGGAGTACGCGGATATCTCAGGAGCATCGGC 576  
 Db 304 AACGACCGCTCCAGCGGAGCTCAAGGAGTGGCTCTCTGGGCTCAAGAGCGGACCTCGGC 363  
 QY 577 ATCGATCGCTGGGCTTCGACTACGTCAGGGGTATGCTCCCTGGGTCTGTCAGGAGCTGG 636  
 Db 364 TTGACGCGTGGGCTTTGACTTCGCCAGGGGTACTCGCGGAGATGCCAAGGTGTAC 423  
 QY 637 CTGAAGTG 644  
 Db 424 ATCGACGG 431

# RESULT 7

CNS0100L/c  
 LOCUS  
 DEFINITION  
 CNS0100L 681 bp DNA linear GSS 14-JUN-2001  
 Anopheles gambiae GSS T7 end of clone 26N13 of Notredame library  
 from strain PEST of Anopheles gambiae (African malaria mosquito),  
 genomic survey sequence.  
 ACCESSION  
 AL153781  
 VERSION  
 AL153781.1 GI:7014700  
 KEYWORDS  
 GSS.  
 SOURCE  
 Anopheles gambiae (African malaria mosquito)

ORGANISM Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
REFERENCE 1 (bases 1 to 681)  
AUTHORS Direct Submission  
TITLE Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
JOURNAL  
REFERENCE 2 (bases 1 to 681)  
AUTHORS Roth, C.W., Brey, P.T., Ke, Z., Collins, F.H. and Weissenbach, J.  
TITLE Direct Submission  
JOURNAL Submitted (16-FEB-2000) BMMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France  
COMMENT This clone is from an A. Gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.  
FEATURES  
source 1..681  
Location/Qualifiers  
/organism="Anopheles gambiae"  
/mol\_type="genomic DNA"  
/strain="PEST"  
/db\_xref="taxon:7165"  
/clone="26N13"  
/clone\_lib="NotreDame1"  
/note="end : T7"  
ORIGIN  
Query Match 4.8%; Score 62.6; DB 29; Length 681;  
Best Local Similarity 54.8%; Pred. No. 0.0019;  
Matches 166; Conservative 0; Mismatches 134; Indels 3; Gaps 2;  
QY 36 CATAAATCAGCGTTCCTACTGGGACGTCCTTCAGGAGGATATGGTGGGACACATACG 95  
Db 302 CTTCTGCAATTTTCACATGTTACTACCCGCGCGGAACTGTGGCAGGAGCGCGC 243  
QY 96 GCAGAGATACCGAGTGTACGATGCGGAATCTCGCAATATGGAATCCCCCGCGAG 155  
Db 242 CGAGCGCGCGCCGACCTGCGCGAATGCGGATTTACGACCTGT-GCTGCGCTGCTA 184  
QY 156 CAAGGATGCGCGCGCTTTCGATGGGTAGGACCCCTAGACTTTCTTACCTCGG 215  
Db 183 CAAGGGCGCTTCGCGCGCTACTCCGCGGTACGATACCTAGATCTGTTCGATCTCGG 124  
QY 216 TGAGTACACACAGAGGAGCGGTAGACGCGCTTTGGGTCCAAAGCAGGAGTCTGTGAA 275  
Db 123 CGAATTTGACCAAGAGGACGCTGGCCCAATACGGGACAGCGCGCTCGAACA 64  
QY 276 --CATGATAACACCGCCACGCTATGGCATGAGTAATAGCCGATATAGTATCAAC 333  
Db 63 CGCTCGGCAACTTCTCGCGGAAACAGCGGTGGGACACTCTACGACGTGGTGTCAAC 4  
QY 334 CAC 336  
Db 3 CAC 1  
RESULT 8  
CG302842 858 bp DNA linear GSS 25-AUG-2003  
LOCUS CG302842  
DEFINITION CG302842 858 bp DNA linear GSS 25-AUG-2003  
genomic survey sequence.  
ACCESSION CG302842  
VERSION CG302842.1 GI:34217056  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 858)

AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Other GSSs: OG2A322TH  
Contact: Cathy Whitelaw  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TP  
Class: sheared ends.  
FEATURES  
source 1..858  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0742C19"  
/clone\_lib="ZM\_0.7-1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb methylation filtered genomic DNA library"  
ORIGIN  
Query Match 4.8%; Score 62.4; DB 29; Length 858;  
Best Local Similarity 47.4%; Pred. No. 0.0023;  
Matches 186; Conservative 0; Mismatches 206; Indels 0; Gaps 0;  
QY 253 GGTCTCAAGCAGGAGCTCGTGAACATGATAACACCGCCCGCTATGGCATGAAGTGA 312  
Db 223 GGCACAGGAGCGCAGCTCAAGTCCCTGATCGAGGCGTTCCACAGCAAGGCGTGAAGTC 282  
QY 313 ATAGCGGATATAGTATCATCAACACCGCGCGGTGACCTGGAGTGAACCCCTTCGTG 372  
Db 283 ATCGCGGACATCGTCAACACCGCGCGGAGCACCAGGCGCGCGGCTCTTAC 342  
QY 373 AAGGACTATACCTGGACCGACTTCTCAAAGGTGCGGTGCGGTAAATACACGCCCAACTAC 432  
Db 343 TGCATCTTCAGGCGCGGACCGCGGACTCCCGCTCGACTGGGCGCCCAATGATCTGC 402  
QY 433 CTCGACTTCCACCGAAGAGCTCCATCGGGCGGATTCGGAAACATTTGGAGGCTATCCC 492  
Db 403 AGCGACGACAGGCGCTACTCCGACGGCACCGGGAACCCCGACACCGGCGCGGCTTCGGC 462  
QY 493 GACATATGCCAGACAGAGCTGGGACGACTGCTCTGGGCGGACGAGAGGCTAC 552  
Db 463 GCGGCGCGGACATCGACCTGAACCGCGCGCTCCAGAGGAGCTCACCGGCTGGCTC 522  
QY 553 GCGGCATATCTCAGGAGCATCGCATCGCTCGCGCTTCGACTAGCTCAAGGGCTAT 612  
Db 523 AACTGGCTCAAGACGAGCGTGGCTTCGACGGGTGGCGCTCGACTTCGCCAAGGCTAC 582  
QY 613 GCTCCCTGGTCTCAGGACTGGCTGAAC 644  
Db 583 TCCGCGGAGCTGCGCAAGGTCTTACATCGACGG 614  
RESULT 9  
CG213284 691 bp DNA linear GSS 22-AUG-2003  
LOCUS CG213284  
DEFINITION CG213284 691 bp DNA linear GSS 22-AUG-2003  
genomic survey sequence.  
ACCESSION CG213284  
VERSION CG213284.1 GI:34113172  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 691)

**AUTHORS** Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

**TITLE** Consortium for Maize Genomics

**JOURNAL** Unpublished (2002)

**COMMENT** Other\_GSSs: OG0EW78TV

Contact: Cathy Whitelaw

**TIGR**

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

**FEATURES** Location/Qualifiers

source

1..519  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0702N11"  
/clone\_lib="ZM 0.7 1.5 kb"  
/note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

**ORIGIN**

Query Match 4.6%; Score 60.8; DB 29; Length 691;  
Best Local Similarity 47.2%; Pred. No. 0.0049;  
Matches 195; Conservative 0; Mismatches 207; Indels 0; Gaps 0;  
QY 253 GGCTCCAGCAGGAGCTCGTAACATGATAAACACCGCCCGCTATGGCATGAAGTA 312  
DB 129 GCGACGAGCGCAGCTCAAGTCCCTGATCGAGGGGTTCCACAGCAAGGCGTGAAGTC 188  
QY 313 ATAGCGGATATAGTATCAACACCGCGCGCGTGTGACTTGGAGTGAACCCCTTCGTG 372  
DB 189 ATCGCGGACATCGTCATCAACACCGCAGGAGGAGCACCAGAGCGCGCGGCATCTAC 248  
QY 373 AACGACTATACCTGGACCGACTTCTCAAGGTGCGGTGGGTAAATACACGCGCAACTAC 432  
DB 249 TGAATCTTCAGGGCGGCGACGCGGACTCCGCTCGACTGGGCGCCCAATGATCTGC 308  
QY 433 CTGACTTCCACCGCAAGAGCTCCATCGCGCGATTCGGGAACATTTGGAGGCTATCCC 492  
DB 309 AGCGACGACAGGCGGTACTCCACGCGCAGCGGGAACCCCGACACCGGCGCGACTTCGCG 368  
QY 493 GACATATGCGACAGAGCTGGGACAGTACTGGCTCTGGCGCAGCAGGAGGCTAC 552  
DB 369 GCGCGCGGAGATCGACCACTGAACCGCGCGTCCAGAGGAGCTCACCGGTGGCTC 428  
QY 553 GCGGCATATCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGACTACGTCAAGGGCTAT 612  
DB 429 AACTGSGCTCAAGACGAGCTGGCTTCGACGCGTGGCGCTCGACTTCGCGCAAGGCTAC 488  
QY 613 GCTCCCTGGGTGTCAGGACTGGCTGACTG 644  
DB 489 TCGCGGACGTCGCCAAGGCTACATCGACGG 520

**RESULT 10**  
CB884205  
LOCUS  
DEFINITION  
5-PRIME, mRNA sequence.  
ACCESSION  
CB884205  
VERSION  
CB884205.1  
KEYWORDS  
EST.  
SOURCE  
Hordeum vulgare subsp. vulgare  
ORGANISM  
Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.  
REFERENCE  
1 (bases 1 to 519)

**AUTHORS** Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner, A.

**TITLE** Barley ESTs from germinating seeds

**JOURNAL** Unpublished (2002)

**COMMENT** Contact: Stein Nils

Molecular Markers Group, Department Genbank

Institute of Plant Genetics and Crop Plant Research (IPK)

Corrensstr. 3, 06466, Gatersleben, Germany

Tel: 039482-5522

Fax: 039482-5595

Email: stein@ipk-gatersleben.de

Insert Length: 519 Std Error: 0.00

Plate: 10 row: M column: 8

Seq primer: M13rev.

**FEATURES** Location/Qualifiers

1..519  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="barke"  
/sub\_species="vulgare"  
/db\_xref="GABI:269134"  
/db\_xref="taxon:112509"  
/clone="HV10M08"  
/tissue\_type="germinating seeds"  
/dev\_stage="germinating seeds (48-96 h)"  
/lab\_host="X110-Gold"  
/clone\_lib="HV"  
/notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Roots were grown for two days on filter paper at room temperature. Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable. Average insert size is 1 kb"

**ORIGIN**

Query Match 4.4%; Score 58; DB 14; Length 519;  
Best Local Similarity 46.7%; Pred. No. 0.019;  
Matches 184; Conservative 0; Mismatches 210; Indels 0; Gaps 0;  
QY 253 GGCTCCAGCAGGAGCTCGTAACATGATAAACACCGCCCGCTATGGCATGAAGTA 312  
DB 106 GGCAACAAGCGCAGCTCAAGTCCCTCATCGGCGGCTCCACGCGCAAGGCGTCAAGGCC 165  
QY 313 ATAGCGGATATAGTATCAACACCGCGCGGTGACTGGAGTGAACCCCTTCGTG 372  
DB 166 ATCGCGCATCTGATCAACACCGCAGCGGAGCGCAAGGACGCGCGGCATCTAC 225  
QY 373 AACGACTATACCTGGACCGACTTCTCAAGGTGCGGTAAATACACGCGCAACTAC 432  
DB 226 TGCATCTTCAGGCGCGCACCCCGGACGCGGCTCGACTGGGCGCCCAATGATCTGC 285  
QY 433 CTGACTTCCACCGCAAGAGCTCCATCGGCGGATTCGGAACATTTGGAGGCTATCC 492  
DB 286 CGCGACGACCGCGCTTACCGGACGCGCACCGGAACCCGACACCGGCGCGGCTTCGGG 345  
QY 493 GACATATGCCACGACAGAGCTGGGACCACTGCTCTGGGCGCAGCAGGAGCTAC 552  
DB 346 GCGCGCGCGGACATCGACCACTCAACCCCGCGCTCCAGAGGAGCTCGTCGAGTGGCTC 405  
QY 553 GCGGCATATCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGACTACGTCAAGGGCTAT 612  
DB 406 AACTGSGCTCAGGACCGACGTCGCTTCGACGCGTTCGACTTCGCCAAGGCTAC 465  
QY 613 GCTCCCTGGGTGTCAGGACTGGCTGAACTGCT 646  
DB 466 TCGCGGACGTCGCCAAGATCTACGTGACCGCT 499

**RESULT 11**



AL820356  
LOCUS AL820356 N:130 Triticum aestivum cDNA clone H07\_N130\_plate\_32, mRNA 422 bp linear EST 15-JUL-2002  
DEFINITION sequence.  
ACCESSION AL820356  
VERSION AL820356.1 GI:21831956  
KEYWORDS EST.  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
1 (bases 1 to 422)  
AUTHORS Wilson, I., Beswick, R., Shepherd, S., Barker, G., Parker, J., Owen, P.,  
Edwards, D., Coghill, J., Holdsworth, M., Lenton, J., Shewry, P. and  
Edwards, K.  
TITLE A BESRC-funded wheat EST resource for the academic community  
JOURNAL Unpublished (2002)  
COMMENT Contact: Barker G  
Institute of Arable Crop Research  
Long Ashton, Bristol BS41 9AF United Kingdom.  
FEATURES  
source Location/Qualifiers  
1..422  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="mercia"  
/db\_xref="taxon:4565"  
/clone="H07\_N130\_plate\_32"  
/tissue\_type="embryo"  
/dev\_stage="2 days post germination"  
/clone\_lib="N:130"  
ORIGIN  
Query Match 4.3%; Score 56.6; DB 9; Length 422;  
Best Local Similarity 47.4%; Pred. No. 0.035;  
Matches 170; Conservative 0; Mismatches 189; Indels 0; Gaps 0;  
QY 253 GGCTCCAGCAGGAGCTCGTGAACATGATAAAACACCGCCACGCTATGGCATGAAGTA 312  
Db 62 GGCACAAAGGCGAGCTCAAGTCCCTGATCGGGCGCTCCACGCAAGGCGGTCAAGGCC 121  
QY 313 ATAGCCATATAGTCATCAACACCGCGCGCGGTGACCTCGAGTGGAAACCCCTTCGTG 372  
Db 122 ATCGCGGACATCGTCATCAACACCGCGCGGAGCGCAAGGACGCGCGGCGCATCTAC 181  
QY 373 AACGACTATACCTGGACGAGCTCTCAAAGGTGCGTGGGTAAATACACGGCCAACTAC 432  
Db 182 TGCATCTTCGAGGCGCGACCCCTGACGCGGCTCGACTGGGCCCCACATGATCTGC 241  
QY 433 CTCGACTTCACCCGAAACGAGCTCCATGCGGGGATTCGGAACATTTGAGGCTATCCC 492  
Db 242 CGCGACGACCGGCCCTACGCGACGCGCAACCCAGACACCGGCGCGGACTTCGGG 301  
QY 493 GACATATGCCACGACAGAGCTGGGACGAGTCTGGCTCTGGCCAGCAGGAGAGCTAC 552  
Db 302 GCGCCCGGACATCGACCTCAACCCGCGGTCCAGAAAGAGTCTGCGAGTGGCTC 361  
QY 553 GCGGCATATCTCAGGAGCATGGCATGCTGCGGCTTCGACTACGTCACGAGGCTA 611  
Db 362 AACTGGCTCAGGACCGAGCTCGGCTTCGACGCTGGGCTTCGACTTCGCAAGGCTA 420

RESULT 12  
CG289556/c  
LOCUS CG289556 593 bp DNA linear GSS 25-AUG-2003  
DEFINITION OGSAR78TC ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0819M12,  
genomic survey sequence.  
ACCESSION CG289556  
VERSION CG289556.1 GI:34203770  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 593)  
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
COMMENT Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.  
FEATURES  
source Location/Qualifiers  
1..593  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0819M12"  
/clone\_lib="ZM\_0.7\_1.5\_KB"  
/notes="Vector: pBSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 4.1%; Score 54; DB 29; Length 593;  
Best Local Similarity 48.2%; Pred. No. 0.16;  
Matches 184; Conservative 0; Mismatches 195; Indels 3; Gaps 1;  
QY 264 GGAGCTCGTGAACATGATAAACACCGCCACGCTATGGCATGAAGTAATAGCGATAT 323  
Db 422 GGAGCTCAAGTCCCTGATCGCGCGTTCACGCGAAGGCGGTGAGGTGGTGGCGACAT 363  
QY 324 AGTCATCAACACCGCGCGGTGACCTGGAGTGGAAACCCCTTCGTGAACGACTATAC 383  
Db 362 GTCATCAACACCGGTGCGCGCACTACAGGACGCGCGGCACTACTGTGATCTTGA 303  
QY 384 CTGGACCGACTTCTCAAAGGTGCGTGGGTAAATACACGGCCAACTACTCGACTTCCA 443  
Db 302 GGGCGGACGCGCGCGCTGGACTGGGGCCCCACATGATCTCGCGGACGACAC 243  
QY 444 CCGGACGAGTCCATGCGGGGATTCGGACATTTGGAGGCTATCCGACATATGCA 503  
Db 242 CATATCTCGGACGCGCGCACTCGACACCGGCGCGACTTCGCGCGGCGCGCA 183  
QY 504 ---CGACAGAGCTGGACCACTGCTGCTGGGCCAGCCAGGAGAGCTACGCGGCATA 560  
Db 182 CATCGACCACTCAACGACCGGTCCAGGGGAGCTCACCGACTGGCTGTGCTCAA 123  
QY 561 TCTCAGGAGCATCGGCATCGATGCTGGCGCTTCGACTACGTCAGGCTATGTCCTG 620  
Db 122 GTCGTCGACCTGGCTTCGACGCTGGCGCTCGACTTCGCAAGGCTACTTCCCGCA 63  
QY 621 GGTCGTCAGGACTGGCTGAAC 642  
Db 62 GGTCCGCAAGGTCTACGTGCAC 41

RESULT 13  
CG345046  
LOCUS CG345046 602 bp DNA linear GSS 26-AUG-2003  
DEFINITION OGVDEL11TH ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0509B21,  
genomic survey sequence.  
ACCESSION CG345046  
VERSION CG345046.1 GI:34262312  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1. (bases 1 to 602)  
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OGVDELITV  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
Location/Qualifiers  
1..602  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0509B21"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 4.1%; Score 54; DB 29; Length 602;  
Best Local Similarity 48.2%; Pred. No. 0.16;  
Matches 184; Conservative 0; Mismatches 195; Indels 3; Gaps 1;

QY 264 GGAGCTCGTGAACATGATTAACACCGCCGCTATGCGATGAAGGTAATAGCGATAT 323  
DB 105 GGAGCTCAAGTCCCTGATCGCCGGTTCACCGCAAGGGCGTGCAGGTGGTGGGACAT 164  
QY 324 AGTCATCAACACCGCGCGGTGACCTGAGTGGAAACCCCTTCGTGAAGACTATAC 383  
DB 165 CGTCATCAACACCGGTGCGCGGACTACAAAGACCGCGGCATCTACTGCTTCGA 224  
QY 384 CTGACCGACTTCTCAAGGTGCGTGGGTAAATACCGGCAACTACCTCGACTTCCA 443  
DB 225 GGGCGGACGCGCGCGCGCTGACCTGGGCCCCACATGATCTGGCGGACGAC 284  
QY 444 CCGAAGCGAGCTCCATCGCGGCGATTCGGAAACATTTGAGGCTATCCGACATNGCA 503  
DB 285 CATATCTCGACGCGGACGCGCAACCTCGACACCGCGCGGACTTCGCGCGGCGCGA 344  
QY 504 ---CGACAAGAGTGGGACGAGTCTGGCTCTGGGCGGAGAGCTACGCGGCATA 560  
DB 345 CATCGACCACTCAACGACGCGGTCCAGCGGAGGTCAACGCTGGCTCTGGCTCAA 404  
QY 561 TCTCAGAGAGTGGGACGAGTCTGGCTCTGGGCGGAGAGCTACGCGGCATA 620  
DB 405 GTGCTCGGACCTGGGCTTCGACGCTGGGCGGCTCGACTTGGCAAGGGCTACTCCCGCA 464  
QY 621 GTGCTCAAGGAGTGGCTGAAC 642  
DB 465 GTGCTCAAGGAGTGGCTGAAC 486

RESULT 14  
CG222587 608 bp DNA linear GSS 22-AUG-2003  
LOCUS OGVF507H ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0565103,  
DEFINITION genomic survey sequence.  
ACCESSION CG222587  
VERSION CG222587.1 GI:34122475  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1. (bases 1 to 608)  
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OGVF50TV  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
Location/Qualifiers  
1..608  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0565103"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBGSK; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 4.1%; Score 54; DB 29; Length 608;  
Best Local Similarity 48.2%; Pred. No. 0.16;  
Matches 184; Conservative 0; Mismatches 195; Indels 3; Gaps 1;

QY 264 GGAGCTCGTGAACATGATTAACACCGCCGCTATGCGATGAAGGTAATAGCGATAT 323  
DB 66 GGAGCTCAAGTCCCTGATCGCCGGTTCACCGCAAGGGCGTGCAGGTGGTGGGACAT 125  
QY 324 AGTCATCAACACCGCGCGGTGACCTGGAGTGGAAACCCCTTCGTGAAGACTATAC 383  
DB 126 CGTCATCAACACCGGTGCGCGGACTACAAAGACCGCGGCGCATCTACTGCTTCGA 185  
QY 384 CTGACCGACTTCTCAAGGTGCGTGGGTAAATACGCGGCAACTACTCGACTTCCA 443  
DB 186 GGGCGGACGCGCGCGCTGACCTGGGCCCCACATGATCTGGCGGACGAC 245  
QY 444 CCGAAGCGAGTCCATCGCGGCGATTCGGAAACATTTGAGGCTATCCGACATNGCA 503  
DB 246 CATATCTCGACGCGGACGCGCAACCTCGACACCGCGCGGACTTCGCGCGGCGCGA 305  
QY 504 ---CGACAAGAGTGGGACGAGTCTGGCTCTGGGCGGAGAGCTACGCGGCATA 560  
DB 306 CATCGACCACTCAACGACGCGGTCCAGCGGAGGTCAACGCTGGCTCTGGCTCAA 365  
QY 561 TCTCAGAGAGTGGGACGAGTCTGGCTCTGGGCGGAGAGCTACGCGGCATA 620  
DB 366 GTGCTCGGACCTGGGCTTCGACGCTGGGCGGCTCGACTTGGCAAGGGCTACTCCCGCA 425  
QY 621 GTGCTCAAGGAGTGGCTGAAC 642  
DB 426 GTGCTCAAGGAGTGGCTGAAC 447

RESULT 15  
CG344571 695 bp DNA linear GSS 16-MAY-2003  
LOCUS OGVF54TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0343J11,  
DEFINITION genomic survey sequence.  
ACCESSION CG344571  
VERSION CG344571.1 GI:30813977  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 695)  
 REFERENCE  
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
 Consortium for Maize Genomics  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Cathy Whitelaw  
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: Tg  
 Class: sheared ends.

FEATURES  
 source  
 1..695  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0343J11"  
 /clone\_lib="ZM 0.7 1.5 KB"  
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

## ORIGIN

Query Match 4.1%; Score 54; DB 28; Length 695;  
 Best Local Similarity 48.2%; Pred.No. 0.17;  
 Matches 184; Conservative 0; Mismatches 195; Indels 3; Gaps 1;  
 QY 264 GGAGCTCGTGAACATGAACACCGCCCGCCCTATGGCATGAGGTAATAGCCGATAT 323  
 Db 192 GGAGCTCAAGTCCCTGATCGCCGGTTCCACGGCAAGGGGTGCAGGTGGTGGGACAT 251  
 QY 324 AGTCATCAACACCGCCCGCGCTGACCTGGAGTGGAAACCCCTTCGTGAACGACTATAC 383  
 Db 252 CGTCATCAACACCGGTGGCGGCTACTACAGGAGCCCGGGCATCTACTGCATCTCGA 311  
 QY 384 CTGACCGACTTCTCAAGTTCGGTGGGTAAATACACGGCCAACTACCTGACTTCCA 443  
 Db 312 GGGCGGCACGCCCGACCGCCCTGGACTGGGGCCCCACATGATCTGCCGCGACGACAC 371  
 QY 444 CCCGAACGAGCTCCATCGGGCGATTCCGGAACATTTGGAGGCTATCCCGACATATGCCA 503  
 Db 372 CATATATCTGGACGGACGGCCACCTCGACACCGGCGCGACTTCGCGCGGGCCCGA 431  
 QY 504 ---CGACAAGAGCTGGGACCACTAGTGGCTCTGGGCCAGCCAGGAGAGCTACCGGGCATA 560  
 Db 432 CATCGACCACTCAACGACCGCTCCAGCGGAGCTCACCGACTGGCTGCTCTGGCTCAA 491  
 QY 561 TCTCAGGAGCATCGGCATCGATGGCTGGCTTCGACTAGTCAAGGGCTATGCTCCCTG 620  
 Db 492 GTCGTCGACCTGGGCTTCGACGCGCTGGCGCTGGACTTCGCCAAGGGCTACTCCCCCGA 551  
 QY 621 GGTCGTCGAAGGACTGGCTGAAC 642  
 Db 552 GGTGCCCAAGGCTACGTCGAC 573

Search completed: June 29, 2004, 11:46:31  
 Job time : 3772 secs



[illegible]

Db 250 GAATTAGAACGAGCGATTGTTGTTAAAGGCAACGGGATTCAAGTGTATGGCGATGTT 309

QY 325 GTCAATCAACCGCGCGCGCGTGA 350

Db 310 GTTATGAACCAATAAAGCGCGAGCTGA 335

## RESULT 2

US-09-540-715A-3  
; Sequence 3, Application US/09540715A  
; Patent No. 6623948  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Borchert, Torben  
; APPLICANT: Nielsen, Bjarne  
; APPLICANT: Nielsen, Vibeke  
; APPLICANT: Hoeck, Lisbeth  
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nucleic A  
; FILE REFERENCE: Encoding Same  
; FILE REFERENCE: 5821.010-US  
; CURRENT APPLICATION NUMBER: US/09/540,715A  
; PRIORITY FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/291,023  
; PRIOR FILING DATE: 2001-08-14  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 3  
; LENGTH: 1458  
; TYPE: DNA  
; ORGANISM: Bacillus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1458)  
US-09-540-715A-3

Query Match 6.1%; Score 79.6; DB 4; Length 1458;  
Best Local Similarity 54.9%; Pred. No. 1.1e-12;  
Matches 179; Conservative 0; Mismatches 144; Indels 3; Gaps 1;  
QY 25 AAGCGCGGGTCAATATGAGCGGTTCTACTGGGACGTCCTTCAAGGAGGAATATGGTGG 84  
Db 13 AGCAAGCGGAACGATTATGCGATTTTGAATGGAACGTTCCGAATGATGGCAACATGG 72  
QY 85 GACACAATACGCGAGAGATACCGGATGCGTACGATCGCGAATCTCCGCAATATGGATT 144  
Db 73 AACCGTTACACACACGCTCAAAATTTAAATAATGCCGAATACAGCAATCTGGATT 132  
QY 145 CCCCCGGGAGCAAGGGCATGGCGCGCGCTATTTCGATGGGCTACGACCCCTACGACTTC 204  
Db 133 CCACCTCGTGAAGAGGACGAGCCAAATATG--TAGGCTACGGTGGCTATGACCTT 189  
QY 205 TTGACCTCGGTAGTACGACAGAGGAGCGGTAGAGCGCTTTGGCTCCAAAGCAG 264  
Db 190 TATGACCTTGGTGAATTTAAACCAAAAGGACGGTCCGTACGAATATGGAAACAAAGCA 249  
QY 265 GAGCTCGTGAACATGATAAACCCGCCATGCGCTATGGCATGAAGGTAAATAGCCGATATA 324  
Db 250 GAATTAGACGAGCGATTGTTGTTAAAGGCAACGGGATTCAGTGTATGGCGATGTT 309  
QY 325 GTCAATCAACCGCGCGCGCGGTGA 350  
Db 310 GTTATGAACCAATAAAGCGCGAGCTGA 335

## RESULT 3

US-08-814-052-5  
; Sequence 5, Application US/08814052  
; Patent No. 6015783  
; GENERAL INFORMATION:  
; APPLICANT: von der Osten, Claus  
; APPLICANT: Cherty, Joel R.  
; APPLICANT: Bjornvad, Mads E.  
; APPLICANT: Vind, Jesper

; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6015783o No. 6015783disk of No. 6015783th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/814,052  
; FILING DATE: 06-MAR-1997  
; CLASSIFICATION: 510  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4684.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELEX:

; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1683 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-814-052-5

Query Match 5.9%; Score 76.8; DB 3; Length 1683;  
Best Local Similarity 54.7%; Pred. No. 7.2e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCAATATGAGCGGTTCTACTGGGACGTCCTTCAAGGAGGAATATGGTGGACACA 90  
Db 100 GGGACGCTGATGCGATTTTGAATGCTATGCTCCCAATGCGGCCCAACATTTGGAAGCGT 159  
QY 91 ATACGCGAGAGATACCGGATGCTGATGCGGAATCTCCGCAATATGGATTTCCTCCCG 150  
Db 160 TTGCAAAACGACTCGGCATATTTCGCTGAACAGGATTTACTGCGCTCTGGATTTCCTCCG 219  
QY 151 GCAGCAAGGCGATGGCGCGCGCTATTTCGATGGGCTACGACCCCTACGACTTCTTTGAC 210  
Db 220 GCATATAAGGAGCAGAGCAAGCGGATG--TGGCTACGGTGTCTTACGACCTTTATGAT 276  
QY 211 CTGGTGTAGTACGACAGAGGAAAGCGGTAGAGACGCTTTGGCTCCAGGAGGCTC 270  
Db 277 TTAGGCGAGTTTCATCAAAAGGCGGTTCCGACAAAGTACGGCAACAAAGAGAGCTG 336  
QY 271 GTGAACATGATAAACACCGCCCGCTATGGCATGAAGGTAAATAGCCGATATAGTTCATC 330  
Db 337 CAATCTGCGATCAAAAGTCTTCATTCCTCCGCGACATTAAAGTTTACGGGATGTGGTTCATC 396  
QY 331 AACCAACCGCGCGCGGTGA 350  
Db 397 AACCAAAAGCGCGCTGA 416

## RESULT 4

US-08-812-829-5  
; Sequence 5, Application US/08812829  
; Patent No. 6017751  
; GENERAL INFORMATION:  
; APPLICANT: von der Osten, Claus  
; APPLICANT: Bjornvad, Mads E.



APPLICANT: Wind, Jesper  
APPLICANT: Rasmussen, Michael Dolberg  
TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING  
TITLE OF INVENTION: OR STAINS FROM CELLULOSIC FABRIC  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 60177510 No. 6017751disk of No. 6017751th America, Inc.  
STREET: 405 Lexington Avenue, Suite 6400  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/812,829  
FILING DATE: 06-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4690.204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1683 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-812-829-5

Query Match 5.9%; Score 76.8; DB 3; Length 1683;  
Best Local Similarity 54.7%; Pred. No. 7.2e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATGACGGGGTCTTACTGGGAGCTGCTTTCAGGAGGAATATGTTGGACACA 90  
Db 100 GGGAGCTGATGCAATATTTTGAATGTACATGCCCAATGACGCGCAATGGAAGCGT 159  
QY 91 ATACGGCAGAGATACGGAGTGTAGTCGCGGAATCTCCGCAATATGATTTCCCGG 150  
Db 160 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCGCTCGGATTCGCGG 219  
QY 151 GCGAGCAGGCGCATGGGCGGCGCTTTCGATGGCTACGACCGCTTCTTTGAC 210  
Db 220 GCATATAGGGAACGAGCCAGCGGATG---TGGGCTACGGTGTACGACCTTTATGAT 276  
QY 211 CTCGGTGAATACGACCAAGGGAACGGTAGAGACGGCTTTGGCTCCAGAGGAGCTC 270  
Db 277 TTAGGGAGTTCATCAAAAAGGAGCGGTTCCGCAAAAGTACGCGCAAAAAGGAGAGCTG 336  
QY 271 GTGACATGATAACACCGCCCGCTATGCGCATGAGTAAATAGCCGATATAGTCATC 330  
Db 337 CAATCTCGGATCAAAAGTCTTCATTCGCGACATTAACGTTTACGGGGATGTGTCATC 396  
QY 331 AACACCGCGCGCGGCTGA 350  
Db 397 AACCAAAAGGCGGCTGA 416

RESULT 5  
US-08-146-422-33  
; Sequence 33, Application US/08146422  
; Patent No. 5543576  
; GENERAL INFORMATION:  
; APPLICANT: VAN OOLJEN, ALBERT J. J.

APPLICANT: RIETVELD, KRIJN  
APPLICANT: HOEKEMA, ANDREAS  
APPLICANT: PEN, JAN  
APPLICANT: SIJMONS, PETER C.  
APPLICANT: VERWOERD, TEUNIS C.  
APPLICANT: QUAX, WILHEMUS J.  
TITLE OF INVENTION: PRODUCTION OF ENZYMES IN SEEDS AND THEIR  
TITLE OF INVENTION: USE  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/146,422  
FILING DATE: 02-NOV-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: KENNEDY, BILL  
REGISTRATION NUMBER: 33,407  
REFERENCE/DOCKET NUMBER: 44615-20011.23  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-146-422-33

Query Match 5.9%; Score 76.8; DB 1; Length 1777;  
Best Local Similarity 54.7%; Pred. No. 7.4e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATGACGGGGTCTTACTGGGAGCGCTTCAGGAGGAATATGTTGGACACA 90  
Db 109 GGGAGCTGATGCAATATTTTGAATGGTACATGCCCAATGACGCGCAATGGAAGCGT 168  
QY 91 ATACGGCAGAGATACCGGAGTGTAGTACGATGCCGAATCTCCGCAATATGATTTCCCGG 150  
Db 169 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCGCTCGGATTCGCGG 228  
QY 151 GCGAGCAGGCGCATGGGCGGCGCTTTCGATGGCTAGGACCGCTTCTTTGAC 210  
Db 229 GCATATAGGGAACGAGCGGATG---TGGGCTACGGTGTACGACCTTTATGAT 285  
QY 211 CTCGGTGAATACGACCAAGGGAACGGTAGAGACGGCTTTGGCTCCAGAGGAGCTC 270  
Db 286 TTAGGGAGTTCATCAAAAAGGAGCGGTTCCGCAAAAGTACGCGCAAAAAGGAGAGCTG 345  
QY 271 GTGACATGATAACACCGCCCGCTATGCGCATGAGGTAATAGCCGATATAGTCATC 330  
Db 346 CAATCTCGGATCAAAAAGTCTTCATTCGCGGACATTAACGTTTACGGGGATGTGTCATC 405  
QY 331 AACACCGCGCGGCTGA 350  
Db 406 AACCAAAAGGCGGCTGA 425

RESULT 6  
US-08-253-575-1  
; Sequence 1, Application US/08253575

Patent No. 5705375  
GENERAL INFORMATION:  
APPLICANT: VAN COYEN, ALBERT J.J.  
APPLICANT: RIETVELD, KRIJN  
APPLICANT: QUAX, WILHELMUS J.  
APPLICANT: VAN DEN ELZEN, PETRUS J.M.  
APPLICANT: PEN, JAN  
APPLICANT: HOEKEMA, ANDREAS  
APPLICANT: SIJMONS, PETER C.  
TITLE OF INVENTION: TRANSGENIC PLANTS HAVING A MODIFIED  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/253,575  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/849,422  
FILING DATE: 12-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20033.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-253-575-1

Query Match 5.9%; Score 76.8; DB 1; Length 1777;  
Best Local Similarity 54.7%; Pred. No. 7.4e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCATATGAGCGGTTCTACTGGGAGCGTCTTCAGGAGGAATATGGTGGGACACA 90  
DB 109 GGGAGCCTGATCGAGTATTTTGAATGGTACATGCCCAATGACGGCCAACTTGGAAAGCGT 168  
QY 91 ATACGGCAGAGATACCGGAGTGGTACGATGCCGGAATCTCCGCAATATGATTCCTCCCG 150  
DB 169 TTGCAAAACGATCGCGCATATTTGGCTGAACACGGTATTACTGCCGCTGGGATTCCTCCCG 228  
QY 151 GCGAGCAGGCGATGGCGCGCTATTTCGATGGGTACGACCCCTACGACTTCTTTGAC 210  
DB 229 GCATATAGGGAGACGACGCAAGCGATG---TGGGCTACGGTCTTACGACCTTTATGAT 285  
QY 211 CTCGGTGTAGTACACGAGAGGAGCGGTAGAGACGGCTTTGGCTTCCAGCAGAGCTC 270  
DB 286 TTAGGGGAGTTTTCATCAAAAGGGAGCGGTTCCGACAAAAGTACGGCACAAGAGAGCTG 345  
QY 271 GTGAACATGATAAACCAGCCCGCTATGCGATGAAGGTAATAGCCGATATAGTCATC 330  
DB 346 CAATCTGCGATCAAAAGTCTTCAATCCCGGACATTAACGTTTACGGGATGTGTCATC 405  
QY 331 AACCAACCGCGCGCGGTGA 350  
|||||

DB 406 AACCAAAAGGGCGGCTGA 425  
RESULT 7  
US-08-626-554-16  
Sequence 16, Application US/08626554  
Patent No. 5714474  
GENERAL INFORMATION:  
APPLICANT: VAN COIJEN, ALBERT J.J.  
APPLICANT: RIETVELD, KRIJN  
APPLICANT: HOEKEMA, ANDREAS  
APPLICANT: PEN, JAN  
APPLICANT: SIJMONS, PETER C.  
APPLICANT: VERWOERD, TEUNIS C.  
TITLE OF INVENTION: PRODUCTION OF ENZYMES IN SEEDS AND THEIR  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 PENNSYLVANIA AVENUE NW  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1888  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/626,554  
FILING DATE: 02-APR-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 26192-20011.10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030 MESNFOERSWASH  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1777 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-626-554-16

Query Match 5.9%; Score 76.8; DB 1; Length 1777;  
Best Local Similarity 54.7%; Pred. No. 7.4e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCATATGAGCGGTTCTACTGGGAGCGTCTTCAGGAGGAATATGGTGGGACACA 90  
DB 109 GGGAGCCTGATCGAGTATTTTGAATGGTACATGCCCAATGACGGCCAACTTGGAAAGCGT 168  
QY 91 ATACGGCAGAGATACCGGAGTGGTACGATGCCGGAATCTCCGCAATATGATTCCTCCCG 150  
DB 169 TTGCAAAACGATCGCGCATATTTGGCTGAACACGGTATTACTGCCGCTGGGATTCCTCCCG 228  
QY 151 GCGAGCAGGCGATGGCGCGCTATTTCGATGGGTACGACCCCTACGACTTCTTTGAC 210  
DB 229 GCATATAGGGAGACGACGCAAGCGATG---TGGGCTACGGTCTTACGACCTTTATGAT 285  
QY 211 CTCGGTGTAGTACACGAGAGGAGCGGTAGAGACGGCTTTGGCTTCCAGCAGAGCTC 270  
DB 286 TTAGGGGAGTTTTCATCAAAAGGGAGCGGTTCCGACAAAAGTACGGCACAAGAGAGCTG 345  
QY 271 GTGAACATGATAAACCAGCCCGCTATGCGATGAAGGTAATAGCCGATATAGTCATC 330  
DB 346 CAATCTGCGATCAAAAGTCTTCAATCCCGGACATTAACGTTTACGGGATGTGTCATC 405

QY 331 AACACCGCGCGCGGTGA 350  
|||||  
Db 406 AACCAAAAGCGCGCTGA 425

## RESULT 8

US-08-468-700-33  
; Sequence 33, Application US/08468700  
; Patent No. 5736499  
; GENERAL INFORMATION:

; APPLICANT: COLIN MITCHINSON  
; APPLICANT: CAROL A. REQUADT  
; APPLICANT: TRACI H. ROPP  
; APPLICANT: LEIF P. SOLHEIM  
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE  
; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS: 40  
; ADDRESSEE: Genencor International

; STREET: 180 Kimball Way  
; CITY: South San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468,700

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 252

; ATTORNEY/AGENT INFORMATION:

; NAME: Stone, Christopher

; REGISTRATION NUMBER: 35,696

; REFERENCE/DOCKET NUMBER: GC275

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 742-7555

; TELEFAX: (415) 742-7217

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1968 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-468-700-33

Query Match 5.9%; Score 76.8; DB 1; Length 1968;

Best Local Similarity 54.7%; Pred. No. 7.7e-12;

Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 31 GGGGTCTAATGCGCGGCTTCTACTGGACGCTGCTTCAGGAGGAATATGGTGGGACACA 90  
|||  
Db 263 GGGACGCTGATGCAGTATTTGAATGGTACATGCTCCCAATGACGCGCAACATTGGAAAGCGT 322

QY 91 ATACGGCAGAGATACCGAGTGGTACGATCCGGAATCTCCGCAATATGATTTCCCGG 150  
|||  
Db 323 TTGCAAAACGACTCGGCATATTGGCTGAACACGGTATTACTGCGCTTGGATTCCCGG 382

QY 151 GCGAGCAAGGCGATGGCGGCGCTTCTCGATGGCTACGACCCCTACGACTTCTTTGAC 210  
|||  
Db 383 GCATATAAGGAAACGAGCAAGCGGATG--TGGCTACGGTGTCTAGGACTTTATGAT 439

QY 211 CTCGGTGTAGTACGACAGAGGAAACGGTAGAGACGCGCTTTGGTCCAAACAGAGCTC 270  
|||  
Db 440 TTAGGGAGTTTTCATCAAAAAGGGACGGTTCCGACAAAGTAGTACGGCACAAGAGAGGTG 499

QY 271 GTGAACATGATAAACACCGCCCTATGCGATGAAGGTAATAGCCGATATAGTCATC 330  
|||  
Db 500 CAATCTCGCATCAAAAGTCTTCATTCGCGGACATTAACGTTTACGGGATGTGTCATC 559

QY 331 AACACCGCGCGCGGTGA 350  
|||||  
Db 560 AACCAAAAGCGCGCTGA 579

## RESULT 9

US-08-645-971-1

; Sequence 1, Application US/08645971

; Patent No. 5763385

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Modified Alpha-Amylases Having Altered

; TITLE OF INVENTION: Calcium Binding Properties

; NUMBER OF SEQUENCES: 5

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/645,971

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1968 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Bacillus licheniformis

US-08-645-971-1

Query Match

Best Local Similarity 54.7%; Score 76.8; DB 1; Length 1968;

Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 31 GGGGTCTAATGCGCGGCTTCTACTGGACGCTGCTTCAGGAGGAATATGGTGGGACACA 90  
|||  
Db 263 GGGACGCTGATGCAGTATTTGAATGGTACATGCTCCCAATGACGCGCAACATTGGAAAGCGT 322

QY 91 ATACGGCAGAGATACCGAGTGGTACGATGCGGAATCTCCGCAATATGATTTCCCGG 150  
|||  
Db 323 TTGCAAAACGACTCGGCATATTGGCTGAACACGGTATTACTGCGCTTGGATTCCCGG 382

QY 151 GCGAGCAAGGCGATGGCGGCGCTTCTCGATGGCTACGACCCCTACGACTTCTTTGAC 210  
|||  
Db 383 GCATATAAGGAAACGAGCAAGCGGATG--TGGGCTACGCTGCTTACGACCTTTATGAT 439

QY 211 CTCGGTGTAGTACGACAGAGGAAACGGTAGAGACGCGCTTTGGCTCCAAACAGAGGCTC 270  
|||  
Db 440 TTAGGGAGTTTTCATCAAAAAGGGACGGTTCCGACAAAGTAGTACGGCACAAGAGAGCTG 499

QY 271 GTGAACATGATAAACACCGCCCTATGCGATGAAGGTAATAGCCGATATAGTCATC 330  
|||  
Db 500 CAATCTCGCATCAAAAGTCTTCATTCGCGGACATTAACGTTTACGGGATGTGTCATC 559

QY 331 AACACCGCGCGCGGTGA 350  
|||||  
Db 560 AACCAAAAGCGCGCTGA 579

## RESULT 10

US-08-468-220-31

; Sequence 31, Application US/08468220

; Patent No. 5824532

; GENERAL INFORMATION:

; APPLICANT: Antrim, Richard L.

; APPLICANT: Barnett, Christopher

; APPLICANT: Mitchinson, Colin

; APPLICANT: Power, Scott D.

; APPLICANT: Requa, Carol

; APPLICANT: Solheim, Leif P.

; TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase

NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,220  
FILING DATE: 06-JUN-95  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/194,664  
FILING DATE: 10-FEB-94  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/016,395  
FILING DATE: 11-FEB-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Stone, Christopher L.  
REGISTRATION NUMBER: 35,696  
REFERENCE/DOCKET NUMBER: GC220D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7555  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1968 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-468-220-31

Query Match 5.9%; Score 76.8; DB 1; Length 1968;  
Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATGACGCGTCTTCTACTGGAGCGTCTTCTAGGAGGAATATGGTGGACACA 90  
DB 263 GGGACGCTGATGCGATATTTTGATGGTACATGCCCAATGACGGCAACATTGGAGCGT 322  
QY 91 ATACGGCAGAGATACCGGAGTGGTACGATGGCGGAATCTCCGCAATATGGATTCCCCCG 150  
DB 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCCGCTCTGGATTCCCCCG 382  
QY 151 GCAGCAGGCGATGGCGCGCTATTTCGATGGCTAGCACCCCTACGACTTCTTTGAC 210  
DB 383 GCATATAGGGAACGAGCAAGCGGATG---TGGGCTACGGTCTTACGACCTTTATGAT 439  
QY 211 CTCGGTGAATACCAAGGGAACCGGTAGACGCGCTTTGGCTTCAAGCAGGAGCTC 270  
DB 440 TTAGGGGAGTTTCATCAAAAAGGACGGTTCGGACAAAGTACGGCAAAAAGGAGGCTG 499  
QY 271 GTGAACATGATAACACCGCCCGCTATGGCATGAGGTATAGCCGATATAGTCATC 330  
DB 500 CAATCTGCGATCAAAAGTCTTCAATCCCGCAATTAACGTTTACGGGATGTGTCATC 559  
QY 331 AACACCGCGCGCGGTGA 350  
DB 560 AACCAAAAGGCGCGCTGA 579

RESULT 11  
US-08-468-698-31  
Sequence 31, Application US/08468698  
Patent No. 5845549  
GENERAL INFORMATION:

APPLICANT: Antrim, Richard L.  
APPLICANT: Barnett, Christopher  
APPLICANT: Mitchinson, Colin  
APPLICANT: Power, Scott D.  
APPLICANT: Requaft, Carol  
APPLICANT: Solheim, Leif P.  
TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,698  
FILING DATE: 06-JUN-95  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/194,664  
FILING DATE: 10-FEB-94  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/016,395  
FILING DATE: 11-FEB-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Stone, Christopher L.  
REGISTRATION NUMBER: 35,696  
REFERENCE/DOCKET NUMBER: GC220D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7555  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1968 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-468-698-31

Query Match 5.9%; Score 76.8; DB 2; Length 1968;  
Best Local Similarity 54.7%; Pred. No. 7.7e-12;  
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;  
QY 31 GGGGTCTAATGACGCGTCTTCTACTGGAGCGTCTTCTAGGAGGAATATGGTGGACACA 90  
DB 263 GGGACGCTGATGCGATATTTTGATGGTACATGCCCAATGACGGCAACATTGGAGCGT 322  
QY 91 ATACGGCAGAGATACCGGAGTGGTACGATGGCGGAATCTCCGCAATATGGATTCCCCCG 150  
DB 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCCGCTCTGGATTCCCCCG 382  
QY 151 GCAGCAGGCGATGGCGCGCTATTTCGATGGCTAGCACCCCTACGACTTCTTTGAC 210  
DB 383 GCATATAGGGAACGAGCAAGCGGATG---TGGGCTACGGTCTTACGACCTTTATGAT 439  
QY 211 CTCGGTGAATACCAAGGGAACCGGTAGACGCGCTTTGGCTTCAAGCAGGAGCTC 270  
DB 440 TTAGGGGAGTTTCATCAAAAAGGACGGTTCGGACAAAGTACGGCAAAAAGGAGGAGCTG 499  
QY 271 GTGAACATGATAACACCGCCCGCTATGGCATGAGGTATAGCCGATATAGTCATC 330  
DB 500 CAATCTGCGATCAAAAGTCTTCAATCCCGCAATTAACGTTTACGGGATGTGTCATC 559  
QY 331 AACACCGCGCGCGGTGA 350  
DB 560 AACCAAAAGGCGCGCTGA 579

## RESULT 12

US-08-704-706A-33  
; Sequence 33, Application US/08704706A  
; Patent No. 5958739  
; GENERAL INFORMATION:  
; APPLICANT: GENENCOR INTERNATIONAL, INC.  
; APPLICANT: COLIN MITCHINSON  
; APPLICANT: ANTHONY G. DAY  
; APPLICANT: ANDREW SHAW  
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International, Inc.  
; STREET: 925 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1013  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,706A  
; FILING DATE: February 20, 1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stone, Christopher  
; REGISTRATION NUMBER: 35,696  
; REFERENCE/DOCKET NUMBER: GC275-3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 846-7555  
; TELEFAX: (650) 845-6405  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1968 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-704-706A-33

Query Match 5.9%; Score 76.8; DB 2; Length 1968;

Best Local Similarity 54.7%; Pred. No. 7.7e-12;

Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY	31	GGGGTCATAATGACGGCGTTCTACTGGGACGTCCTTCAGGAGGAATATGTTGGGACACA	90
DB	263	GGGACGCTGATGCATATTTTGAATGGTACATGCCCAATGACGCCCAACATTGGAAGCGT	322
QY	91	ATACGGCAGAAGATACCGGAGTGGTACGATCCGGAATCTCCGCAATATGATTCCCGCG	150
DB	323	TTGCAAAACGACTCGCGCATATTTGGCTGAACACGGTATTACTGCCGCTTGGATTCCCGCG	382
QY	151	GGGACCAAGGCGATGGCGGCGCTTATTCGATGGCTACGACCCCTTACGACTTCTTTGAC	210
DB	383	GCATATAAGGGAACGAGCCACCGGATG---TGGGCTACGGTCTTACGCTTTATGAT	439
QY	211	CTCGGTGAGTACGACAGAAGGAAACGGTAGAGACGCGCTTTGGCTCCAAAGAGAGCTC	270
DB	440	TTAGGGGAGTTTCATCAAAAAGGGACGGTTCCGACAAAGTACGGCACAAGAGAGAGCTG	499
QY	271	GTGAACATGATTAACACCGCCACCGCTATGGCATGAAGGTAAATAGCCGATATAGTCATC	330
DB	500	CAATCTGCGATCAAAAGTCTTCATTTCCCGGACATTTAAGCTTTACGGGGATGTGGTCATC	559
QY	331	AACCAACCGCGCGCGGTGA	350
DB	560	AACCAAAAGGCGCGCTGA	579

## RESULT 14

US-08-914-679A-1  
; Sequence 1, Application US/08914679A

## RESULT 13

US-08-890-383-1  
; Sequence 1, Application US/08890383  
; Patent No. 6008026  
; GENERAL INFORMATION:  
; APPLICANT: Anthony G. Day  
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE HAVING INTRODUCED  
; THEREIN A DISULFIDE BOND  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International, Inc.  
; STREET: 925 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1013  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/890,383  
; FILING DATE: To Be Assigned  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stone, Christopher  
; REGISTRATION NUMBER: 35,696  
; REFERENCE/DOCKET NUMBER: GC377  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 846-7555  
; TELEFAX: (650) 845-6504  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1968 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-890-383-1

Query Match 5.9%; Score 76.8; DB 3; Length 1968;

Best Local Similarity 54.7%; Pred. No. 7.7e-12;

Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY	31	GGGGTCATAATGACGGCGTTCTACTGGGACGTCCTTCAGGAGGAATATGTTGGGACACA	90
DB	263	GGGACGCTGATGCATATTTTGAATGGTACATGCCCAATGACGCCCAACATTGGAAGCGT	322
QY	91	ATACGGCAGAAGATACCGGAGTGGTACGATCCGGAATCTCCGCAATATGATTCCCGCG	150
DB	323	TTGCAAAACGACTCGCGCATATTTGGCTGAACACGGTATTACTGCCGCTTGGATTCCCGCG	382
QY	151	GGGACCAAGGCGATGGCGGCGCTTATTCGATGGCTACGACCCCTTACGACTTCTTTGAC	210
DB	383	GCATATAAGGGAACGAGCCACCGGATG---TGGGCTACGGTCTTACGCTTTATGAT	439
QY	211	CTCGGTGAGTACGACAGAAGGAAACGGTAGAGACGCGCTTTGGCTCCAAAGAGAGCTC	270
DB	440	TTAGGGGAGTTTCATCAAAAAGGGACGGTTCCGACAAAGTACGGCACAAGAGAGAGCTG	499
QY	271	GTGAACATGATTAACACCGCCACCGCTATGGCATGAAGGTAAATAGCCGATATAGTCATC	330
DB	500	CAATCTGCGATCAAAAGTCTTCATTTCCCGGACATTTAAGCTTTACGGGGATGTGGTCATC	559
QY	331	AACCAACCGCGCGCGGTGA	350
DB	560	AACCAAAAGGCGCGCTGA	579

```

; Patent No. 6080568
; GENERAL INFORMATION:
; APPLICANT: Anthony G. Day
; APPLICANT: Barbara A. Swanson
; TITLE OF INVENTION: MUTANT ALPHA-AMYLASE COMPRISING MODIFICATION
; FILE REFERENCE: AT RESIDUES CORRESPONDING TO A210, H405 AND/OR
; TITLE OF INVENTION: T412 IN BACILLUS LICHENIFORMIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/914,679A
; FILING DATE: To Be Assigned
; ATTORNEY/AGENT INFORMATION:
; NAME: Stone, Christopher
; REGISTRATION NUMBER: 35,696
; REFERENCE/DOCKET NUMBER: GC387
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 846-7555
; TELEFAX: (650) 945-6504
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-914-679A-1

Query Match 5.9%; Score 76.8; DB 3; Length 1968;
Best Local Similarity 54.7%; Pred. No. 7.7e-12;
Matches 175; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 31 GGGGTCAATATGACGGGTTCTACTGGGACGTGCTTCTAGGAGGAATATGTTGGACACA 90
Db 263 GGGACGCTGATGTCAGTATTTTGAATGGTACATGCCCAATGACGGCCAAACATTTGGAAGCGT 322
QY 91 ATACGGCAGAAAGATACCGGAGTGTACGATGCGGGAATCTCCGCAATATGATTCCTCCCG 150
Db 323 TTGCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCCGTCTGGATTCCCGCG 382
QY 151 GCGAGCAAGGGCATGGCGCGGCTATTTCGATGGGCTACGACCCCTACGACTTCTTTGAC 210
Db 383 GCATATAAGGGAACGAGCCAAAGCGGATG---TGGGCTACGGTCTTACGACCTTTATGAT 439
QY 211 CTCGGTGAGTACGACGAGGGAACGGTAGACGCGCTTTGGCTCCAAAGCAGGAGCTC 270
Db 440 TTAGGGGAGTTTCATCAAAAGGAGCGTTTCGGACAAAGTACGGCACAAAGGAGAGCTG 499
QY 271 GTGAACATGATAAACACCGCCCGCTATGTCATGAAGGTAATAGCCGATATAGTCATC 330
Db 500 CAATCTCGCATCAAAAGTCTTCATTCCCGGACATTAAACGTTTACGGGGATGTTGTCATC 559
QY 331 AACCAACCGCCCGCGGTGA 350
Db 560 AACCAAAAGGCGCGCTGA 579

Search completed: June 29, 2004, 11:48:28
Job time : 116 secs.

```

```

RESULT 15
US-08-985-659-33
; Sequence 33, Application US/08985659
; Patent No. 621134
; GENERAL INFORMATION:

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 10:43:46 ; Search time 607 Seconds  
(without alignments)  
9929.476 Million cell updates/sec

Title: US-10-081-739A-1

Perfect score: 1311

Sequence: 1 atggccaagtactcgagct.....gctactgcgggtgggtctga 1311

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3031105 seqs, 2298700234 residues

Total number of hits satisfying chosen parameters: 6062210

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US04\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1311	100.0	1311	15	US-10-146-662-1
2	1311	100.0	1311	15	US-10-105-733-1
3	1311	100.0	1311	15	US-10-081-739A-1
4	1239	94.5	1311	15	US-10-081-872-5
5	1239	94.5	1311	16	US-10-385-305-5
6	1231	93.9	1311	15	US-10-081-872-71
7	1231	93.9	1311	16	US-10-385-305-71
8	1192.6	91.0	1311	15	US-10-081-872-37
9	1192.6	91.0	1311	16	US-10-385-305-37
10	1186.2	90.5	1311	15	US-10-081-872-39
11	1186.2	90.5	1311	16	US-10-385-305-39
12	1183	90.2	1311	15	US-10-081-872-29
13	1183	90.2	1311	16	US-10-385-305-29
14	1171.8	89.4	1311	15	US-10-081-872-13

15	1171.8	89.4	1311	16	US-10-385-305-13	Sequence 13, Appl
16	1160.6	88.5	1311	15	US-10-081-872-61	Sequence 61, Appl
17	1160.6	88.5	1311	16	US-10-385-305-61	Sequence 61, Appl
18	1155.8	88.2	1311	15	US-10-081-872-55	Sequence 55, Appl
19	1155.8	88.2	1311	16	US-10-385-305-55	Sequence 55, Appl
20	1147.8	87.6	1311	15	US-10-081-872-3	Sequence 3, Appl
21	1147.8	87.6	1311	16	US-10-081-872-25	Sequence 25, Appl
22	1147.8	87.6	1311	15	US-10-081-872-45	Sequence 45, Appl
23	1147.8	87.6	1311	16	US-10-385-305-3	Sequence 3, Appl
24	1147.8	87.6	1311	16	US-10-385-305-25	Sequence 25, Appl
25	1147.8	87.6	1311	16	US-10-385-305-45	Sequence 45, Appl
26	1146.2	87.4	1311	15	US-10-081-872-59	Sequence 59, Appl
27	1146.2	87.4	1311	16	US-10-385-305-59	Sequence 59, Appl
28	1144.6	87.3	1311	15	US-10-081-872-51	Sequence 51, Appl
29	1144.6	87.3	1311	16	US-10-385-305-51	Sequence 51, Appl
30	1143	87.2	1311	15	US-10-081-872-63	Sequence 63, Appl
31	1143	87.2	1311	16	US-10-385-305-63	Sequence 63, Appl
32	1141.4	87.1	1311	15	US-10-081-872-19	Sequence 19, Appl
33	1141.4	87.1	1311	16	US-10-385-305-19	Sequence 19, Appl
34	1139.8	86.9	1311	15	US-10-081-872-15	Sequence 15, Appl
35	1139.8	86.9	1311	16	US-10-385-305-15	Sequence 15, Appl
36	1138.2	86.8	1311	15	US-10-081-872-11	Sequence 11, Appl
37	1138.2	86.8	1311	16	US-10-385-305-11	Sequence 11, Appl
38	1130.2	86.2	1311	15	US-10-081-872-41	Sequence 41, Appl
39	1130.2	86.2	1311	16	US-10-385-305-41	Sequence 41, Appl
40	1120.6	85.5	1311	15	US-10-081-872-1	Sequence 1, Appl
41	1120.6	85.5	1311	16	US-10-385-305-1	Sequence 1, Appl
42	1119	85.4	1311	15	US-10-081-872-35	Sequence 35, Appl
43	1119	85.4	1311	16	US-10-385-305-35	Sequence 35, Appl
44	1117.4	85.2	1311	15	US-10-081-872-23	Sequence 23, Appl
45	1117.4	85.2	1311	16	US-10-081-872-31	Sequence 31, Appl

#### ALIGNMENTS

RESULT 1  
US-10-146-662-1  
; Sequence 1, Application US/10146662  
; Publication No. US2003003172A1  
; GENERAL INFORMATION:  
; APPLICANT: Genendash, Joel  
; TITLE OF INVENTION: NOVEL METHODS OF ENZYME PURIFICATION  
; FILE REFERENCE: 09010-109001  
; CURRENT APPLICATION NUMBER: US/10/146,662  
; CURRENT FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1308)  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-146-662-1

Query Match	100.0%;	Score 1311;	DB 15;	Length 1311;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1311;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	ATGCCCAAGTACTCCGAGCTGGAAAAGGGGGGTCATATGCAGCGCTTCTACTCGGAC	60	
Db	1	ATGCCCAAGTACTCCGAGCTGGAAAAGGGGGGTCATATGCAGCGCTTCTACTCGGAC	60	
QY	61	GTGCTTCAGGAGGAATATGGTGGGACACAATACGCCAGAAGATACCGAGTGGTACGAT	120	
Db	61	GTGCTTCAGGAGGAATATGGTGGGACACAATACGCCAGAAGATACCGAGTGGTACGAT	120	
QY	121	GCCGGAATCTCCGCAATATGGATTCCCCCGGCGAGGACATGGCGCGCCTATTTCG	180	

Db 121 GCGGGAATCTCCGCAATATGGATTTCCCGCGGAGCAAGGGCATATGGCGCGCCCTATTTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGAACGGTA 240  
Qy 241 GAGACGGCTTTGGCTCCAGAGGAGCTCGTAACATATAAACAACCGCCACCGCTAT 300  
Db 241 GAGACGGCTTTGGCTCCAGAGGAGCTCGTAACATATAAACAACCGCCACCGCTAT 300  
Qy 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGAGGTGG 360  
Db 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGAGGTGG 360  
Qy 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTGCGCTCGGTTAAATAC 420  
Db 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTGCGCTCGGTTAAATAC 420  
Qy 421 ACGGCAACTACCTGACTTCCACCGAGAGCTCCATGCGGCGGATTCCGGAACATTT 480  
Db 421 ACGGCAACTACCTGACTTCCACCGAGAGCTCCATGCGGCGGATTCCGGAACATTT 480  
Qy 481 GGAGGCTATCCCGACATATGCGACCAAGAGCTGGACGACTTGGCTCTGGGCGAGC 540  
Db 481 GGAGGCTATCCCGACATATGCGACCAAGAGCTGGACGACTTGGCTCTGGGCGAGC 540  
Qy 541 CAGGAGCTACCGGCAATCTCAGGAGCTCGGATCGATCGCTGGGCGCTTCGACTAC 600  
Db 541 CAGGAGCTACCGGCAATCTCAGGAGCTCGGATCGATCGCTGGGCGCTTCGACTAC 600  
Qy 601 GTCAAGGCTATGCTCCCTGGTCTGCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660  
Db 601 GTCAAGGCTATGCTCCCTGGTCTGCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660  
Qy 661 GTTGGAGAGTACGGGACCAACGCTGACGCTGTTCTCAACTGGGCGATCTCGAGCGT 720  
Db 661 GTTGGAGAGTACGGGACCAACGCTGACGCTGTTCTCAACTGGGCGATCTCGAGCGT 720  
Qy 721 GCCAAGTCTTTGACTTGGCCCTCTACTCAAGATGGATGAGGCTTTCACCAACAAAC 780  
Db 721 GCCAAGTCTTTGACTTGGCCCTCTACTCAAGATGGATGAGGCTTTCACCAACAAAC 780  
Qy 781 ATTCAGCGCTGCTCTGCTGCTTCAGACGGCGCAGACTGTTCTCCGCGACCGCTTC 840  
Db 781 ATTCAGCGCTGCTCTGCTGCTTCAGACGGCGCAGACTGTTCTCCGCGACCGCTTC 840  
Qy 841 AAGCGCTAACCTTTGAGCAACCCAGCACCGATATATCTGGAACAGTATCCAGCC 900  
Db 841 AAGCGCTAACCTTTGAGCAACCCAGCACCGATATATCTGGAACAGTATCCAGCC 900  
Qy 901 TACCGCTTCACTCCCTACCTACGAGGCGCCAGCACAATATTTACCGGACTACGAGGAG 960  
Db 901 TACCGCTTCACTCCCTACCTACGAGGCGCCAGCACAATATTTACCGGACTACGAGGAG 960  
Qy 961 TGGCTCAACAGGATAGCTCAAGAACCTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020  
Db 961 TGGCTCAACAGGATAGCTCAAGAACCTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020  
Qy 1021 AGCACCGACATAGTCTACTACGATAAGGAGTCACTCATCTTCGTCAGGAACGGCTACGGG 1080  
Db 1021 AGCACCGACATAGTCTACTACGATAAGGAGTCACTCATCTTCGTCAGGAACGGCTACGGG 1080  
Qy 1081 GACAAGCGGGGCTTATAA CTTACATCAACTAGGCTCGAGCAAGGCGGAGAGGTGGTT 1140  
Db 1081 GACAAGCGGGGCTTATAA CTTACATCAACTAGGCTCGAGCAAGGCGGAGAGGTGGTT 1140  
Qy 1141 TATCTCGGAGTTCGGGGCGCTGATCCACGAGTATCTGTAACCTTCGAGGCTGG 1200  
Db 1141 TATCTCGGAGTTCGGGGCGCTGATCCACGAGTATCTGTAACCTTCGAGGCTGG 1200  
Qy 1201 GTAGACAAGTACGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

Db 1201 GTAGACAAGTACGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Qy 1261 GCCAACGGCAGTATGGCTACTCGGTGAGCTACTTGGGGTGGCTGA 1311  
Db 1261 GCCAACGGCAGTATGGCTACTCGGTGAGCTACTTGGGGTGGCTGA 1311  
RESULT 2  
US-10-105-733-1  
; Sequence 1, Application US/10105733  
; Publication No. US20030138786A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Miller, Carl  
; APPLICANT: Kazaoka, Martin  
; APPLICANT: Short, Jay  
; APPLICANT: Mathur, Eric  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE REFERENCE: 09010-107002  
; CURRENT APPLICATION NUMBER: US/10105,733  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 10/081,739  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated  
US-10-105-733-1  
Query Match 100.0%; Score 1311; DB 15; Length 1311;  
Best Local Similarity 100.0%; Pred. No 0;  
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCCAAGTACTCCGAGCTGGAAAGGGGGGTCTATAATCGAGGGCTTCTACTGGGAC 60  
Db 1 ATGCCAAGTACTCCGAGCTGGAAAGGGGGGTCTATAATCGAGGGCTTCTACTGGGAC 60  
Qy 61 GTGCCCTCAGGAGGAATATGTTGGGACACATACGGCAGAGATACCGGAGTGTACGAT 120  
Db 61 GTGCCCTCAGGAGGAATATGTTGGGACACATACGGCAGAGATACCGGAGTGTACGAT 120  
Qy 121 GCCGGAATCTCCGCAATATGGATTTCCCGCGGAGCAAGGGCATATGGCGCGCCCTATTTCG 180  
Db 121 GCCGGAATCTCCGCAATATGGATTTCCCGCGGAGCAAGGGCATATGGCGCGCCCTATTTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTCTTTTGACCTCGGTGAGTACGACCAAGAGGAACGGTA 240  
Qy 241 GAGACGGCTTTGGCTCCAGAGGAGCTCGTAACATATAAACAACCGCCACCGCTAT 300  
Db 241 GAGACGGCTTTGGCTCCAGAGGAGCTCGTAACATATAAACAACCGCCACCGCTAT 300  
Qy 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGAGGTGG 360  
Db 301 GGCATGAAGTAATAGCCGATATAGTCATCAACACCGCGCGCGGTGACCTGAGGTGG 360  
Qy 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTGCGGTTCGAGTAAATAC 420  
Db 361 AACCCCTTGGTGAACGACTATACCTGGACGACTTCTCAAGGTGCGGTTCGAGTAAATAC 420

```
QY 421 ACGGCAACTACTCGACTTCCACCGACAGAGCTCCATCGGGGATTCGGAAACATTT 480
Db 421 ACGGCAACTACTCGACTTCCACCGACAGAGCTCCATCGGGGATTCGGAAACATTT 480
QY 481 GGAGGCTATCCCGACATATGCCACACAGAGCTGGGACCACTGCTCTGGGCCAGC 540
Db 481 GGAGGCTATCCCGACATATGCCACACAGAGCTGGGACCACTGCTCTGGGCCAGC 540
QY 541 CAGGAGAGTACCGCGCATATCTCAGGAGCATTCGGCATCGATCGCTGGCGTTCGACTAC 600
Db 541 CAGGAGAGTACCGCGCATATCTCAGGAGCATTCGGCATCGATCGCTGGCGTTCGACTAC 600
QY 601 GTCAAGGGCTATGCTCCCTCGGTCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
Db 601 GTCAAGGGCTATGCTCCCTCGGTCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
QY 661 GTTGAGAGTACTGGGACACCAACCTGACAGCTGGCTGAACTGGTGGGAGGCTGGGCG 720
Db 661 GTTGAGAGTACTGGGACACCAACCTGACAGCTGGCTGAACTGGTGGGAGGCTGGGCG 720
QY 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGAGCTGGATGAGGCTTTGACAAACAAAC 780
Db 721 GCCAAGGCTTTGACTTCGCCCTCTACTACAGAGCTGGATGAGGCTTTGACAAACAAAC 780
QY 781 ATTCCAGGCTCGTCTCTGCCCTTCAGAACCGCCAGACTGTTGCTCCCGCGACCCGTTTC 840
Db 781 ATTCCAGGCTCGTCTCTGCCCTTCAGAACCGCCAGACTGTTGCTCCCGCGACCCGTTTC 840
QY 841 AAGGCGCTAACTTTGTAGCAACACACGACCGCATATAAATCTGGAAACAGATTCAGGCC 900
Db 841 AAGGCGCTAACTTTGTAGCAACACACGACCGCATATAAATCTGGAAACAGATTCAGGCC 900
QY 901 TACGGCTTCATCTCACCCTACGAGGGCCAGCCGACATATTTCTACCGGACTACGAGGAG 960
Db 901 TACGGCTTCATCTCACCCTACGAGGGCCAGCCGACATATTTCTACCGGACTACGAGGAG 960
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGCGAGGA 1020
Db 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGCGAGGA 1020
QY 1021 AGCACCGCATAGTCTACTACGATAACGATCACTCATCTTCGTCAGGACGGTACGGG 1080
Db 1021 AGCACCGCATAGTCTACTACGATAACGATCACTCATCTTCGTCAGGACGGTACGGG 1080
QY 1081 GACAAAGCGGGCTTTATAACCTACATCAACCTAGGCTCGAGAACGGCCGGAAGTGGGTT 1140
Db 1081 GACAAAGCGGGCTTTATAACCTACATCAACCTAGGCTCGAGAACGGCCGGAAGTGGGTT 1140
QY 1141 TATGTGCGGAGTTTCGGGGCGGTGCATCCACAGATATCTGTAACCTCGGAGGCTGG 1200
Db 1141 TATGTGCGGAGTTTCGGGGCGGTGCATCCACAGATATCTGTAACCTCGGAGGCTGG 1200
QY 1201 GTAGACAAAGTACGTCTACTCAAGGGGCTGGGTCTATCTCGAAGCTCCAGGCTTACGACCT 1260
Db 1201 GTAGACAAAGTACGTCTACTCAAGGGGCTGGGTCTATCTCGAAGCTCCAGGCTTACGACCT 1260
QY 1261 GCCAACGGGCGATGAGTCTCGTGTGAGGCTACTCGGGGTGGGCTGA 1311
Db 1261 GCCAACGGGCGATGAGTCTCGTGTGAGGCTACTCGGGGTGGGCTGA 1311
```

## RESULT 3

```
US-10-081-739A-1
; Sequence 1, Application US/10081739A
; Publication No. US20030170634A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Miller, Carl
; APPLICANT: Kazaoka, Martin
; APPLICANT: Short, Jay
```

```
; APPLICANT: Mathur, Eric
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; FILE REFERENCE: 09010-107001
; CURRENT APPLICATION NUMBER: US/10/081,739A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated
US-10-081-739A-1
```

```
Query Match 100.0%; Score 1311; DB 15; Length 1311;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCAAGTACTCCGAGCTGGAAAGGGCGGGTTCATATGAGCGCTTCTACTGGGAC 60
Db 1 ATGCCCAAGTACTCCGAGCTGGAAAGGGCGGGTTCATATGAGCGCTTCTACTGGGAC 60
QY 61 GTCCCTTCAGGAGGAATATGTTGGGACACAATACGGCAGAAAGATACCGGAGTGGTACGAT 120
Db 61 GTCCCTTCAGGAGGAATATGTTGGGACACAATACGGCAGAAAGATACCGGAGTGGTACGAT 120
QY 121 GCCGGAATCTCCGCAATATGGATTCCTCCGCGGAGCAAGGGCATGGGCGGCTATTCG 180
Db 121 GCCGGAATCTCCGCAATATGGATTCCTCCGCGGAGCAAGGGCATGGGCGGCTATTCG 180
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGAGGAAACGGTA 240
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGAGGAAACGGTA 240
QY 241 GAGACCGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTAT 300
Db 241 GAGACCGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAACAACCGCCACGCTAT 300
QY 301 GGATCAAGGTAATAGCGCATATAGTCATCAACCAACCGCGCGGCGGTGACCTGAGTGG 360
Db 301 GGATCAAGGTAATAGCGCATATAGTCATCAACCAACCGCGCGGCGGTGACCTGAGTGG 360
QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTGCGGTGCGGTAATAC 420
Db 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTGCGGTGCGGTAATAC 420
QY 421 ACGGCAACTACTCGACTTCCACCGGAACGAGCTCCATGCGGGCGATTCGGGAAACATTT 480
Db 421 ACGGCAACTACTCGACTTCCACCGGAACGAGCTCCATGCGGGCGATTCGGGAAACATTT 480
QY 481 GGAGGCTATCCCGACATATGCCACACAGAGCTGGGACCACTGCTGGGCCAGC 540
Db 481 GGAGGCTATCCCGACATATGCCACACAGAGCTGGGACCACTGCTGGGCCAGC 540
QY 541 CAGGAGAGTACCGCGCATATCTCAGGAGCATTCGAGGAGCTCGATGCTGGCGTTCGACTAC 600
Db 541 CAGGAGAGTACCGCGCATATCTCAGGAGCATTCGAGGAGCTCGATGCTGGCGTTCGACTAC 600
QY 601 GTCAAGGGCTATGCTCCCTCGGTCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
Db 601 GTCAAGGGCTATGCTCCCTCGGTCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
QY 661 GTTGAGAGTACTGGGACACCAACCTGACAGCTGGCTGTTCTCAACTGGGCGATTCGAGCGGT 720
Db 661 GTTGAGAGTACTGGGACACCAACCTGACAGCTGGCTGTTCTCAACTGGGCGATTCGAGCGGT 720
```

; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-5

[illegible]

Db 961 TGCTCAACAAAGGACAGGTTGAACAACTCATATGATACAGCAACCTCGAGGTGA 1020  
Qy 1021 AGCACGACATAGTCTACTACGATAGGATGAATCATCTTGTAGGAAACGGCTACGGG 1080  
Db 1021 AGCACGACATAGTCTACTACGATAGGATGAATCATCTTGTAGGAAACGGCTACGGG 1080  
Qy 1081 GACAAGCCGGGCTTAACTACATCAACCTAGGCTCGAGCAAGCCCGGAAAGGTGGGT 1140  
Db 1081 GACAAGCCGGGCTTAACTACATCAACCTAGGCTCGAGCAAGCCCGGAAAGGTGGGT 1140  
Qy 1141 TATGTCGCAAGTTCGGGCGGCTGATCCAGAGTATCTGTTAACTCGAGAGGTGGG 1200  
Db 1141 TATGTCGCAAGTTCGGGCGGCTGATCCAGAGTATCTGTTAACTCGAGAGGTGGG 1200  
Qy 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Db 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Qy 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGAGCTACTCGGGGTGGGTGA 1311  
Db 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGAGCTACTCGGGGTGGGTGA 1311

RESULT 5

US-10-385-305-5  
; Sequence 5, Application US/10385305  
; Publication No. US20040018607A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Prev, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupska, Malgorzata  
; TITLE OF INVENTION: ENZYMS HAVING ALPHA AMYLASE ACTIVITY  
; FILE OF INVENTION: AND METHODS OF USE THEREOF  
; CURRENT APPLICATION NUMBER: US/10/385,305  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide

Query Match 94.5%; Score 1239; DB 16; Length 1311;  
Best Local Similarity 96.6%; Pred. No. 0;  
Matches 1266; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1 ATGGCCAAAGTACTCCGAGCTGGAAGCGGCGGTCTAATGCGAGGCTTCTACTGGGAC 60  
Db 1 ATGGCCAAAGTACTCCGAGCTGGAAGCGGCGGTCTAATGCGAGGCTTCTACTGGGAC 60  
Qy 61 GTGCGCTTCAGGAGGAATATGTTGGGACACAAATACGCGAGAGATACCGGAGTGTACGAT 120  
Db 61 GTCCAGGTGAGGAATCTGTGGGACACCAATCAGGAGCAGATACCGGAGTGTACGAG 120  
Qy 121 GCCGGAATCTCCGAATATGATTTCCCGCGGAGAGGAGCATGCGCGGCGCTATTCG 180  
Db 121 GCCGGAATCTCTACTCAAGCGGCTGGGTCTATCTCGAGCTCCAGCTTACGACCT 1260

Db 121 CGGGGAATATCCGCCATTTGGATTCCCGCGGCAAGCAAGGCAATGGGCGGCGCTATTTCG 180  
Qy 181 ATGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGACCGGTA 240  
Db 181 ATGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGGACCGGTA 240  
Qy 241 GAGAGCGGCTTTGGCTTCAAGCAGAGCTCGTGAACATGATAACACCGCCACGCTTAT 300  
Db 241 GAGAGCGGCTTTGGCTTCAAGCAGAGCTCGTGAACATGATAACACCGCCACGCTTAT 300  
Qy 301 GGCATGAAGGTAAATAGCGATATAGTTCATCAACCCGCGCGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGGTAAATAGCGATATAGTTCATCAACCCGCGCGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTCGTGAACGACTATACCTGACCGACTTCTCAAAGGTGCGGTAAATATAC 420  
Db 361 AACCCCTTCGTGAACGACTATACCTGACCGACTTCTCAAAGGTGCGGTAAATATAC 420  
Qy 421 ACGGCCAACTACCTCGACTTCCACCCGAAAGCTCCATCGGGCGGATTCGGGAACATTT 480  
Db 421 ACGGCCAACTACCTCGACTTCCACCCGAAAGCTCCATCGGGCGGATTCGGGAACATTT 480  
Qy 481 GGAGCTATCCGACATATGCTCCGACGAGCTGGGACCAAGTACTGGCTCTTGGGCGCAG 540  
Db 481 GGAGCTATCCGACATATGCTCCGACGAGCTGGGACCAAGTACTGGCTCTTGGGCGCAG 540  
Qy 541 CAGGAGAGCTACCGGCAATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAC 600  
Db 541 CAGGAGAGCTACCGGCAATATCTCAGGAGCATCGGATCGATCGCTGGCGCTTCGACTAC 600  
Qy 601 GTCAAGGGCTATGCTCCCTGGTCTGCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660  
Db 601 GTCAAGGGCTATGCTCCCTGGTCTGCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660  
Qy 661 GTTGGAGAGTACTGGGACCAACGCTGAGCTGTTCTCAACTGGGCATATCTCGAGCGGT 720  
Db 661 GTTGGAGAGTACTGGGACCAACGCTGAGCTGTTCTCAACTGGGCATATCTCGAGCGGT 720  
Qy 721 GCCAAGGTCTTTCGACTTGGCCCTTACTACAGATGATGAGGCTTTCGACAAACAAAC 780  
Db 721 GCCAAGGTCTTTCGACTTGGCCCTTACTACAGATGATGAGGCTTTCGACAAACAAAC 780  
Qy 781 ATTCAGCGCTCGTCTGCTCCCTTCAGAACCGGCAAGCTGTTGTCTCCCGGACCGCTTC 840  
Db 781 ATTCAGCGCTCGTCTGCTCCCTTCAGAACCGGCAAGCTGTTGTCTCCCGGACCGCTTC 840  
Qy 841 AAGGCGCTAACCTTCTAGCAACCAAGCAACCGATATATCTGGAACAAAGTATCCAGCC 900  
Db 841 AAGGCGCTAACCTTCTAGCAACCAAGCAACCGATATATCTGGAACAAAGTATCCAGCC 900  
Qy 901 TACGCGTTTCATCTCCTACCTACGAGGCGGCAAGCAATATTTCTACCGGACTACGAGGAG 960  
Db 901 TATGCTTTTCATCTCCTACCTACGAGGCGGCAAGCAATATTTCTACCGGACTACGAGGAG 960  
Qy 961 TGGCTCAACAGGATTAAGCTCAAGAACCTCATCTGATACATGAGAACCTCGCGGAGGA 1020  
Db 961 TGGCTCAACAGGATTAAGCTCAAGAACCTCATCTGATACATGAGAACCTCGCGGAGGA 1020  
Qy 1021 AGCACGACATAGTCTACTACGATAGGATGAATCATCTTGTGAGGAAACGGCTACGGG 1080  
Db 1021 AGCACGACATAGTCTACTACGATAGGATGAATCATCTTGTGAGGAAACGGCTACGGG 1080  
Qy 1081 GACAAGCCGGGCTTAACTACATCAACCTAGGCTCGAGCAAGCCCGGAAAGGTGGGT 1140  
Db 1081 GACAAGCCGGGCTTAACTACATCAACCTAGGCTCGAGCAAGCCCGGAAAGGTGGGT 1140  
Qy 1141 TATGTCGCAAGTTCGGGCGGCTGATCCAGAGTATCTGTTAACTCGAGAGGTGGG 1200  
Db 1141 TATGTCGCAAGTTCGGGCGGCTGATCCAGAGTATCTGTTAACTCGAGAGGTGGG 1200  
Qy 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Db 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAGCTCCAGCTTACGACCT 1260

QY 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTGCGGGTGGGCTGA 1311  
 DB 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTGCGGGTGGGCTGA 1311

## RESULT 6

US-10-081-872-71  
 ; Sequence 71, Application US/10081872  
 ; Publication No. US20030125534A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT APPLICATION NUMBER: US/10/081.872  
 ; CURRENT FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 71  
 ; LENGTH: 1311  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated oligonucleotide  
 US-10-081-872-71

Query Match 93.9%; Score 1231; DB 15; Length 1311;  
 Best Local Similarity 96.2%; Pred. No. 0;  
 Matches 1261; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 ATGGCCAAAGTACTCGAGCTGGAAGAGGGGGGTCAATATGACGGCTTCTACTGGGAC 60  
 DB 1 ATGGCCAAAGTACTCGAGCTGGAAGAGGGGGGTCAATATGACGGCTTCTACTGGGAC 60

QY 61 GTGCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120  
 DB 61 GTGCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120

QY 121 GCCGGAATCTCCGCAATATGATATCCCGGGGAGCAAGGCAATGGCGCGCCTATTGG 180  
 DB 121 GCCGGAATCTCCGCAATATGATATCCCGGGGAGCAAGGCAATGGCGCGCCTATTGG 180

QY 181 ATGGGCTACGACCCCTAGCACTCTTTGACCTGGGTGAGTACGACAGAGGAGCGGTA 240  
 DB 181 ATGGGCTACGACCCCTAGCACTCTTTGACCTGGGTGAGTACGACAGAGGAGCGGTA 240

QY 241 GAGACGGCTTTGGCTCCAGCAGAGAGCTCGTGAAATGATATAACACCGCCACGCGTAT 300  
 DB 241 GAGACGGCTTTGGCTCCAGCAGAGAGCTCGTGAAATGATATAACACCGCCACGCGTAT 300

QY 301 GCGATGAAGGTAATAGCGGATATAGTATCATCAACCGCGCGCGGTGACCTGGAGTGG 360  
 DB 301 GCGATGAAGGTCATAGCGGATATAGTATCATCAACCGCGCGCGGTGACCTGGAGTGG 360

QY 361 AACCCCTTCGGAACGACTATPACTGGACCGCACTTCTCAAAGGTGCGGTGAGGTAATAC 420  
 DB 361 AACCCCTTCGTTGGGGACTACCTGGACCGCACTTCTCAAAGGTGAGGTAATAT 420

QY 421 ACGGCCAACTACTCTCGACTTCCACCCGACGAGCTCCATGTCGGGGGATTCGCGAATTT 480

DB 421 ACTGCCAACTACTCTCGACTTCCACCCGACGAGCTCCATCGCGGGCATTCGGAACATTT 480  
 QY 481 GGAGGCTATCCGACATATATGCCAGACAGAGCTGGGACAGTACTGGCTTGGGGCCAGC 540  
 DB 481 GGAGGCTATCCGACATATATGCCAGACAGAGCTGGGACAGTACTGGCTTGGGGCCAGC 540

QY 541 CAGGAGAGCTACCGGCAATATCTCAGGAGCATCGGCATCGATCGCTGGGCTTCGACTAC 600  
 DB 541 CAGGAGAGCTACCGGCAATATCTCAGGAGCATCGGCATCGATCGCTGGGCTTCGACTAC 600

QY 601 GTCAAGGGCTATGCTCCCTGGGTCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660  
 DB 601 GTCAAGGGCTATGCTCCCTGGGTCTCAAGGACTGGCTGAACTGGTGGGAGGCTGGGG 660

QY 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCATATCTCGAGCGGT 720  
 DB 661 GTTGGAGAGTACTGGGACACCAACGTCGACGCTGTCTCAACTGGGCATATCTCGAGCGGT 720

QY 721 GCCAAGGTCTTTGACTTCGGCTCTACTAACAAGTGGATGAGGCTTTGACAAACAAAC 780  
 DB 721 GCCAAGGTCTTTGACTTCGGCTCTACTAACAAGTGGATGAGGCTTTGACAAACAAAC 780

QY 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAACGGCCAGACTGTTGTCTCCCGCGACCCGCTTC 840  
 DB 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAACGGCCAGACTGTTGTCTCCCGCGACCCGCTTC 840

QY 841 AAGCCGCTAACCTTTGTAGCAACACGACGATATTAATCTGGAAACAAGTATCCAGCC 900  
 DB 841 AAGCCGCTAACCTTTGTAGCAACACGACGATATTAATCTGGAAACAAGTATCCAGCC 900

QY 901 TAGCGGTTTCATCTCTCACTACGAGGCGCAGCGCAATAATCTTACCGGACTTACGAGGAG 960  
 DB 901 TAGCGGTTTCATCTCTCACTACGAGGCGCAGCGCAATAATCTTACCGGACTTACGAGGAG 960

QY 961 TGGCTCAACAAAGGATAGCTCAAGAACCTCTCTGATATGATGAGAACCTCCCGGAGGA 1020  
 DB 961 TGGCTCAACAAAGGATAGCTCAAGAACCTCTCTGATATGATGAGAACCTCCCGGAGGA 1020

QY 1021 AGCACCGCATAGTCTACTACGATACGATGAACCTCATCTCGTCAGGAACGGCTACCGG 1080  
 DB 1021 AGCACCGCATAGTCTACTACGATACGATGAACCTCATCTCGTCAGGAACGGCTACCGG 1080

QY 1081 GACAAGCGGGGCTTATAACCTTACATCAACCTTAGGCTCGAGCAAGCGCGGAGGAGTGGT 1140  
 DB 1081 GACAAGCGGGGCTTATAACCTTACATCAACCTTAGGCTCGAGCAAGCGCGGAGGAGTGGT 1140

QY 1141 TATGTGCCGAGGTTCCGGCGCGTGCATCCACGAGTATCTGTTAACTCGAGGCTGG 1200  
 DB 1141 TATGTGCCGAGGTTCCGGCGCGTGCATCCACGAGTATCTGTTAACTCGAGGCTGG 1200

QY 1201 GTAGACAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
 DB 1201 GTAGACAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

QY 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTGCGGGGTGGGCTGA 1311  
 DB 1261 GCCAACGGGAGTATGGCTACTCCGTGTGGAGTACTGCGGGGTGGGCTGA 1311

## RESULT 7

US-10-385-305-71  
 ; Sequence 71, Application US/10385305  
 ; Publication No. US20040018607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata



; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated oligonucleotide
US-10-385-305-71

Query Match 93.9%; Score 1231; DB 16; Length 1311;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1261; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1 ATGGCCAGTACTCGAGTGGAAAGCGGGGTCTAATATCGAGGCTTCTACTGGGAC 60
DB 1 ATGGCCAGTACTCGAGTGGAAAGCGGGGTCTAATATCGAGGCTTCTACTGGGAC 60
QY 61 GTGCGCTTCAGGAGGAATATGGTGGGACACAATACGCGGAGAGATACCGAGTGGTACGAT 120
DB 61 GTGCGCTTCAGGAGGAATATGGTGGGACACAATACGCGGAGAGATACCGAGTGGTACGAT 120
QY 121 GCGGGAATTCGCGAATATGGATCCCGCGGAGCAAGGCGATGGCGGCGCTTATTCG 180
DB 121 GCGGGAATTCGCGAATATGGATCCCGCGGAGCAAGGCGATGGCGGCGCTTATTCG 180
QY 181 ATGGGCTAGACCCCTACGACTTCTTACCTCGGTAGTACGACCAAGGAGCAACGTA 240
DB 181 ATGGGCTAGACCCCTACGACTTCTTACCTCGGTAGTACGACCAAGGAGCAACGTA 240
QY 241 GAGACGCGCTTTGGTCCAAAGCAGGAGCTCGTGAACATGATAAACCACCGCCACGCTAT 300
DB 241 GAGACGCGCTTTGGTCCAAAGCAGGAGCTCGTGAACATGATAAACCACCGCCACGCTAT 300
QY 301 GGCATGAAGTAAATAGCGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360
DB 301 GGCATGAAGTAAATAGCGGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360
QY 361 AACCCCTTCGTGAACGACTATACCTTGGACCGACTTCTCAAGGTCGGTAAATAC 420
DB 361 AACCCCTTCGTGGGACTACCTTGGACCGACTTCTCAAGGTCGGTAAATAC 420
QY 421 ACGGCAACTACTCGACTTCACCCGAGCAGCTTCCATGCGGCGATTCGGGAACATTT 480
DB 421 ACTGCAACTACTCGACTTCACCCGAGCAGCTTCCATGCGGCGATTCGGGAACATTT 480
QY 481 GGAGGCTATCCGACATATCCAGCAGAGAGCTGGGACGAGTACTGGCTTGGGCGAGC 540
DB 481 GGAGGCTATCCGACATATCCAGCAGAGAGCTGGGACGAGTACTGGCTTGGGCGAGC 540
QY 541 CAGGAGAGCTACGGGCAATATCTCAGGAGCATCGGCAATCGATGCGTGGCGCTTCGATAC 600
DB 541 CAGGAGAGCTACGGGCAATATCTCAGGAGCATCGGCAATCGATGCGTGGCGCTTCGATAC 600
QY 601 GTCAGGCTATGTCCTCGGTGCTCAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
DB 601 GTCAGGCTATGTCCTCGGTGCTCAGGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
QY 661 GTTGAGAGTACTGGGACCAACGTCGACGCTCTCTCAACTGGGCATATCTGAGCGGT 720

DB 661 GTTGAGAGTACTGGGACCAACGTCGACGCTCTCTCAACTGGGCATATCTGAGCGGT 720
QY 721 GCCAAGTCTTTGACTTTCGCTCTACTACAGATGATGAGGCTTTGACAAACAAAAC 780
DB 721 GCCAAGTCTTTGACTTTCGCTCTACTACAGATGATGAGGCTTTGACAAACAAAAC 780
QY 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAAAGCGGCGAGACTGTTGCTCCCGGACCGGTC 840
DB 781 ATTCAGCGCTCGTCTCTGCGCTTCAGAAAGCGGCGAGACTGTTGCTCCCGGACCGGTC 840
QY 841 AAGCGCTAACCTTTGTAGCAACACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900
DB 841 AAGCGCTAACCTTTGTAGCAACACGACACCGATATAATCTGGAACAAAGTATCCAGCC 900
QY 901 TACGCTTTCATCTCCTACCTACGAGGCGGCGAGCAATATCTACCGGACTACGAGGAG 960
DB 901 TACGCTTTCATCTCCTACCTACGAGGCGGCGAGCAATATCTACCGGACTACGAGGAG 960
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTTCGCGGAGGA 1020
DB 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTTCGCGGAGGA 1020
QY 1021 AGCACCACATAGTCTACTACGATACGATGAACTCATCTCTCAGGAAGCGGTACGGG 1080
DB 1021 AGCACCACATAGTCTACTACGATACGATGAACTCATCTCTCAGGAAGCGGTACGGG 1080
QY 1081 GACAAGCGCGGCTTATAAATCACTACATCAACCTTAGGCTCGGCAAGCGGAGAGTGGGT 1140
DB 1081 GACAAGCGCGGCTTATAAATCACTACATCAACCTTAGGCTCGGCAAGCGGAGAGTGGGT 1140
QY 1141 TATGTCGCGAAGTTCGCGGCGGCTGATCCAGAGTATCTGTTAACTTCGAGAAACGCTG 1200
DB 1141 TATGTCGCGAAGTTCGCGGCGGCTGATCCAGAGTATCTGTTAACTTCGAGAAACGCTG 1200
QY 1201 GTAGAACAGTACCTTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
DB 1201 GTAGAACAGTACCTTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
QY 1261 GCCAACGGGCGAGTATGGCTACTCCGTGTGAGACTACTGCGGGTGGGCTGA 1311
DB 1261 GCCAACGGGCGAGTATGGCTACTCCGTGTGAGACTACTGCGGGTGGGCTGA 1311

RESULT 8
US-10-081-872-37
; Sequence 37, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kerovuo, Janne S.
; APPLICANT: Slupski, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081,872
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence



FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-37

Query Match 91.0%; Score 1192.6; DB 15; Length 1311;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 1237; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

QY 1 ATGGCCAGTACTCCGAGTGGGACAAAGGGCGGGGTCTAATATGACGGGTTCTTACTGGGAC 60
DB 1 ATGGCCAAAGTACCTGGAGCTCGAAGAGGGCGGGGTCTAATATGACGGGTTCTTACTGGGAC 60
QY 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120
DB 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACGGCAGAGATACCGGAGTGGTACGAT 120
QY 121 GCGGAAATCTCGGCAATATGATTTCCCGCGGAGAGCGATGGGCGCGCTATTTCG 180
DB 121 GCGGAAATCTCGGCAATATGATTTCCCGCGGAGAGCGATGGGCGCGCTATTTCG 180
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAAACGGTA 240
DB 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGAAACGGTA 240
QY 241 GAGAGCGGTTTGGCTCAACGAGAGGCTCGTGAACATGATTAACACCGCCACCGCTAT 300
DB 241 GAGAGCGGTTTGGCTCAACGAGAGGCTCGTGAACATGATTAACACCGCCACCGCTAT 300
QY 301 GCGATGAAGGTAATAGCCGATATAGTCAATCAACCGCGCGGCGGTGACCTGGAGTGG 360
DB 301 GCGATGAAGGTAATAGCCGATATAGTCAATCAACCGCGCGGCGGTGACCTGGAGTGG 360
QY 361 AACCCCTTCTGTAAGCACTATACCTGAGACGACTTCTCAAGGTGCGGTGGGTAATATAC 420
DB 361 AACCCCTTCTGTAAGCACTATACCTGAGACGACTTCTCAAGGTGCGGTGGGTAATATAC 420
QY 421 ACGGCAACTACTCGACTTCCACCCGAAACGAGTCCATCGCGGCGGATTCGCGAAATTT 480
DB 421 ACGGCAACTACTCGACTTCCACCCGAAACGAGTCCATCGCGGCGGATTCGCGAAATTT 480
QY 481 GGAGGCTATCCGCAATATGACGACGAGCTGGGACCACTGCTGGTCTGGGCGAGC 540
DB 481 GGAGGCTATCCGCAATATGACGACGAGCTGGGACCACTGCTGGTCTGGGCGAGC 540
QY 541 CAGGAGAGCTACGGGCAATATCTCAGGAGATCGGATCGATCGCTGGGCTTCGACTAC 600
DB 541 CAGGAGAGCTACGGGCAATATCTCAGGAGATCGGATCGATCGCTGGGCTTCGACTAC 600
QY 601 GTCAAGGCTATGCTCCCTGGGTCGTCAAGGACTGGCTGAATCTGGTGGGAGGCTGGGG 660
DB 601 GTCAAGGCTATGCTCCCTGGGTCGTCAAGGACTGGCTGAATCTGGTGGGAGGCTGGGG 660
QY 661 GTTGAGAGTACTGGGACACCAAGTCGACGCTGTTCTCAACTGGGCATCTCGAGCGGT 720
DB 661 GTTGAGAGTACTGGGACACCAAGTCGACGCTGTTCTCAACTGGGCATCTCGAGCGGT 720
QY 721 GCCAAGGCTTTGATTTGGCCCTCTACTCAAGATGATGAGGCTTTTGACACAAAAAC 780
DB 721 GCCAAGGCTTTGATTTGGCCCTCTACTCAAGATGATGAGGCTTTTGACACAAAAAC 780
QY 781 ATTCAGGCGCTCGTCTCTGCGCTTCAGAACGGCAGACTGTTCTTCGCGGACCGCTTC 840
DB 781 ATTCAGGCGCTCGTCTCTGCGCTTCAGAACGGCAGACTGTTCTTCGCGGACCGCTTC 840
QY 841 AAGGCGGTAACTTTGTAGCAAAACCAACGACATGATTAATCTGGAAACAGTATCCAGCC 900
DB 841 AAGGCGGTAACTTTGTAGCAAAACCAACGACATGATTAATCTGGAAACAGTATCCAGCC 900
QY 901 TACGCTTCACTCTCACTACGAGGGCCAGCACAATATTTCTACCGGACTACGAGGAG 960
DB 901 TACGCTTCACTCTCACTACGAGGGCCAGCACAATATTTCTATCGGACTACGAGGAG 960
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTTCGCGGAGGA 1020

```

```

DB 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGACAACTCCCGGAGGA 1020
QY 1021 AGCACCGACATAGTCTACTACGATAACGATCACTCATCTTCTCAGGAACGGCTACGGG 1080
DB 1021 AGCACTGACATCGTTTACTACGACAGCAGGAGTCTATTTCTGAGAAACGGCTACGGA 1080
QY 1081 GACAAAGCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGSCCGGAAGGTGGTT 1140
DB 1081 AGCAAGCGGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGSCCGGAAGGTGGTT 1140
QY 1141 TATGTCCGAAGTTCGGGGCGGCTGCTCCAGAGTACTTGTAACTCGAGAGCTGG 1200
DB 1141 TACGTTCCGAAGTTCGGGGCGGCTGCTCCAGAGTACTTGTAACTCGAGAGCTGG 1200
QY 1201 GTAGACAAAGTACGCTACTCAAGCGGCTGGGCTTATCTCGAAGCTCCAGCTTACGACCT 1260
DB 1201 GTGACAAAGTGGGTGGACTCAAGCGGCTGGGCTTATCTCGAGGCTCTCTGCCACGACCG 1260
QY 1261 GCCAACGGGACAGTATGGCTACTCCGTGTGGAGTACTTCGGGGGTGGGCTGA 1311
DB 1261 GCCAACGGGACAGTACGGCTACTCCGTGTGGAGTACTTCGGGGGTGGGCTGA 1311

```

## RESULT 9

US-10-385-305-37  
; Sequence 37, Application US/10385305  
; Publication No. US20040018607A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupska, Malgorzata  
; TITLE OF INVENTION: ENZYMS HAVING ALPHA AMYLASE ACTIVITY  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/385,305  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-385-305-37

Query Match 91.0%; Score 1192.6; DB 16; Length 1311;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 1237; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

QY 1 ATGGCCAAAGTACTCCGAGTGGGACAAAGGGCGGGGTCTAATATGACGGGTTCTTACTGGGAC 60
DB 1 ATGGCCAAAGTACTCCGAGTGGGACAAAGGGCGGGGTCTAATATGACGGGTTCTTACTGGGAC 60
QY 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACCGCAGAGATACCGGAGTGGTACGAT 120
DB 61 GTGGCTTCAGAGGAATATGGTGGGACACAAATACCGCAGAGATACCGGAGTGGTACGAT 120
QY 121 GCGGAAATCTCGGCAATATGATTTCCCGCGGAGAGCGATGGGCGCGCTATTTCG 180

```

Db 121 GCGGAATCTCCGAATATGATTTCCCGCGGAGCAAGGGCATGGCGCGCTATTTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTCTTTTGAACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTCTTTTGAACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Qy 241 GAGACGGCTTTGGCTTCAAGCAGGAGCTGCTGAAACATGATATAACACCGCCCAACGCTAT 300  
Db 241 GAGACGGCTTTGGCTTCAAGCAGGAGCTGCTGAAACATGATATAACACCGCCCAACGCTAT 300  
Qy 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTGTGAACGACTATCTCGGACCGACTTCTCAAGGTGCGGTGAGTAAATAC 420  
Db 361 AACCCCTTGTGAACGACTATCTCGGACCGACTTCTCAAGGTGCGGTGAGTAAATAC 420  
Qy 421 ACGGCAACTACCTCGACTTCCACCGACGAGCTCCATGCGGGGATTTCCGGAACATTT 480  
Db 421 ACGGCAACTACCTCGACTTCCACCGACGAGCTCCATGCGGGGATTTCCGGAACATTT 480  
Qy 481 GAGGCTATCCGACATATGCCAAGAGCTGGACCACTACTGGCTCTGGGCCAGC 540  
Db 481 GAGGCTATCCGACATATGCCAAGAGCTGGACCACTACTGGCTCTGGGCCAGC 540  
Qy 541 CAGGAGACTCGGCGCATATCTCAGGACATCGGATCGCTCTGGCGCTTCGACTAC 600  
Db 541 CAGGAGACTCGGCGCATATCTCAGGACATCGGATCGCTCTGGCGCTTCGACTAC 600  
Qy 601 GTCAAGGGCTATGCTCCCTCGGTGCTCAAGGACTCGCTGAACTGTGGGGAGCTGGCG 660  
Db 601 GTCAAGGGCTATGCTCCCTCGGTGCTCAAGGACTCGCTGAACTGTGGGGAGCTGGCG 660  
Qy 661 GTTGAGAGTACTGGGACCAACCTGCGACGTGTCTCAACTGGGCACTCTCGAGCGGT 720  
Db 661 GTTGAGAGTACTGGGACCAACCTGCGACGTGTCTCAACTGGGCACTCTCGAGCGGT 720  
Qy 721 GCAAGGCTTTGACTTGGCTCTGCTTACTACAGATGATGAGGCTTTGACAAACAAAC 780  
Db 721 GCAAGGCTTTGACTTGGCTCTGCTTACTACAGATGATGAGGCTTTGACAAACAAAC 780  
Qy 781 ATTCAGGCTGCTCTGCTGCTTCAAGAGCGGACAGCTGTGCTCCCGGACCGGTT 840  
Db 781 ATTCAGGCTGCTCTGCTGCTTCAAGAGCGGACAGCTGTGCTCCCGGACCGGTT 840  
Qy 841 AAGGCGTAACCTTTGAGCAACCAACGACGACGATATCTGGAACAAGTATCCAGCC 900  
Db 841 AAGGCGTAACCTTTGAGCAACCAACGACGACGATATCTGGAACAAGTATCCAGCC 900  
Qy 901 TACGGGTTTATCTCCTCAGGAGCGGCGGCAATATTTCTACCGGACTACGAGGAG 960  
Db 901 TACGGGTTTATCTCCTCAGGAGCGGCGGCAATATTTCTACCGGACTACGAGGAG 960  
Qy 961 TGGCTCAACAGGATAGCTCAAGAACTCTATCTGGATACATGAGAACTCCCGGAGGA 1020  
Db 961 TGGCTCAACAGGATAGCTCAAGAACTCTATCTGGATACATGAGAACTCCCGGAGGA 1020  
Qy 1021 AGCACCGCATAGTCTACTAGATAAGATGAACTCATCTTCGTCAGGAACGGTACGG 1080  
Db 1021 AGCACCGCATAGTCTACTAGATAAGATGAACTCATCTTCGTCAGGAACGGTACGG 1080  
Qy 1081 GACAAAGCGGGCTTATAACCTATCACTAGGCTCGAGCAAGCGGGAAGGTGGGTT 1140  
Db 1081 GACAAAGCGGGCTTATAACCTATCACTAGGCTCGAGCAAGCGGGAAGGTGGGTT 1140  
Qy 1141 TATGTCCGAAGTTTCGGGGCGGCTGATCCACGAGTATCTGTAACCTCGGAGGCTGG 1200  
Db 1141 TATGTCCGAAGTTTCGGGGCGGCTGATCCACGAGTATCTGTAACCTCGGAGGCTGG 1200  
Qy 1201 CTAGCAAGTACGCTACTCAAGCGGCTGGGCTCTATCTCGAAGCTCCAGCTTACGACCT 1260

Db 1201 GTGGACAAAGTGGGTGGACTCAAGCGGCTGGTCTACCTCGAGGCTCTCTGCCACGACCCG 1260  
Qy 1261 GCCAACGGGAGTATGGCTACTCTCGTGTGGAGTACTCTCGGGGTGGGTGA 1311  
Db 1261 GCCAACGGGAGTATGGCTACTCTCGTGTGGAGTACTCTCGGGGTGGGTGA 1311  
RESULT 10  
US-10-081-872-39  
; Sequence 39, Application US/10081872  
; Publication No: US20030125534A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupska, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-39  
Query Match 90.5%; Score 1186.2; DB 15; Length 1311;  
Best Local Similarity 94.1%; Pred. No. 0;  
Matches 1233; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
Qy 1 ATGGCAAGTACTCCGAGCTGGAAAGGGCGGGTCAATATGAGCGGCTTCTACTGGGAC 60  
Db 1 ATGGCAAGTACTCCGAGCTGGAAAGGGCGGGTCAATATGAGCGGCTTCTACTGGGAC 60  
Qy 61 GTCCCTTCAGGAGGATATGTTGGGACAAATACGCGCAAGATACCGAGTGGTACGAT 120  
Db 61 GTCCCTTCAGGAGGATATGTTGGGACAAATACGCGCAAGATACCGAGTGGTACGAT 120  
Qy 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGGCTATTTCG 180  
Db 121 GCCGGAATCTCCGCAATATGGATTCCCGCGGAGCAAGGGCATGGCGGCTATTTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Qy 241 GAGACGGCTTTGGCTTCCAAAGCAGGAGCTGCTGAAACATGATATAACACCGCCCAACGCTAT 300  
Db 241 GAGACGGCTTTGGCTTCCAAAGCAGGAGCTGCTGAAACATGATATAACACCGCCCAACGCTAT 300  
Qy 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGGTAATAGCCGATATAGTATCATCAACACCGCGCGCGGTGACCTGGAGTGG 360  
Qy 361 AACCCCTTGTGAACGACTATCTCGGACCGACTTCTCAAGGTGCGGTGAGTAAATAC 420  
Db 361 AACCCCTTGTGAACGACTATCTCGGACCGACTTCTCAAGGTGCGGTGAGTAAATAC 420

QY 421 ACAGGCAACTACCTCGACTTCCACCCGAAAGAGCTCCATCGGGCGGATTCGGAAACATTT 480  
DB 421 ACAGGCAACTACCTCGACTTCCACCCGAAAGAGCTCCATCGGGCGGATTCGGAAACATTT 480  
QY 481 GGAGGCTATCCCGACATATGACAGCAAGAGCTGGGACAGACTGCTGCTCGGGCCAGC 540  
DB 481 GGAGGCTATCCCGACATATGACAGCAAGAGCTGGGACAGACTGCTGCTCGGGCCAGC 540  
QY 541 CAGGAGCTACCGGGCATATCTCAGAGGATTCGGGATCGATCGCTGGCGCTTCGACTAC 600  
DB 541 CAGGAGCTACCGGGCATATCTCAGAGGATTCGGGATCGATCGCTGGCGCTTCGACTAC 600  
QY 601 GTCAAGGCTATGCTCCCTGGGTCTGCTCAAGGAGCTGGGATGCTGGGAGGCTGGGG 660  
DB 601 GTCAAGGCTATGCTCCCTGGGTCTGCTCAAGGAGCTGGGATGCTGGGAGGCTGGGG 660  
QY 661 GTTGAGAGTACTGGGACCAACAGTCCGCTGCTCAAGGAGCTGGGATGCTGGGAGGCT 720  
DB 661 GTTGAGAGTACTGGGACCAACAGTCCGCTGCTCAAGGAGCTGGGATGCTGGGAGGCT 720  
QY 721 GCCAAGGCTTTCGACTTCCGCTCTACTACAGATGATGAGGCTTTGACAAACAAAC 780  
DB 721 GCCAAGGCTTTCGACTTCCGCTCTACTACAGATGATGAGGCTTTGACAAACAAAC 780  
QY 781 ATTCAGGCTGCTGCTTCCGCTCTCAGAGCGGCGAGCTGCTGCTCCGCGACCGCTTC 840  
DB 781 ATTCAGGCTGCTGCTTCCGCTCTCAGAGCGGCGAGCTGCTGCTCCGCGACCGCTTC 840  
QY 841 AAGGCGGTAACTTGTAGCAAAACACAGACACCATATATCTGGAAACAGTATCCAGCC 900  
DB 841 AAGGCGGTAACTTGTAGCAAAACACAGACACCATATATCTGGAAACAGTATCCAGCC 900  
QY 901 TAGCGGCTTATCTCAGGAGCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 901 TAGCGGCTTATCTCAGGAGCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
QY 961 TGGCTCAACAGGATTAAGCTCAAGAACTCATCTGGATATATGAGAACTTCGCGGAGGA 1020  
DB 961 TGGCTCAACAGGATTAAGCTCAAGAACTCATCTGGATATATGAGAACTTCGCGGAGGA 1020  
QY 1021 AGCAGCGATAGTCTACGATAAAGAGTAACTATCTGTCAGGAAACGCTACGGG 1080  
DB 1021 AGCAGCGATAGTCTACGATAAAGAGTAACTATCTGTCAGGAAACGCTACGGG 1080  
QY 1081 GACAGCGCGGCTTATAACCTCATCAACCTAGGCTCGAGGAGGCGGAGGCTGGGT 1140  
DB 1081 GACAGCGCGGCTTATAACCTCATCAACCTAGGCTCGAGGAGGCGGAGGCTGGGT 1140  
QY 1141 TATGTGCGGAAAGTTCGCGGCGGCTGCAATCAAGAGTATCTGGTAACTTCGAGGCTGG 1200  
DB 1141 TATGTGCGGAAAGTTCGCGGCGGCTGCAATCAAGAGTATCTGGTAACTTCGAGGCTGG 1200  
QY 1201 GTAGCAAGTACTGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTTCAGACCTTC 1260  
DB 1201 GTAGCAAGTACTGCTACTCAAGCGGCTGGGTCTATCTCGAAGCTTCAGACCTTC 1260  
QY 1261 GCCAAGCGGAGTATGGCTACTCTCGTGGAGCTATCTGGGGGTGGGCTGA 1311  
DB 1261 GCCAAGCGGAGTATGGCTACTCTCGTGGAGCTATCTGGGGGTGGGCTGA 1311

## RESULT 11

US-10-385-305-39  
; Sequence 39, Application US/10385305  
; Publication No. US20040018607A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.

; APPLICANT: Slupski, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/385,305  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-385-305-39

Query Match 90.5%; Score 1186.2; DB 16; Length 1311;  
Best Local Similarity 94.1%; Pred. No. 0;  
Matches 1233; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 1 ATGCCCAAGTACTCCGAGCTCGGAAAGGCGGGTCTAATGACGCGCTTCTACTGGGAC 60  
DB 1 ATGCCCAAGTACTCCGAGCTCGGAAAGGCGGGTCTAATGACGCGCTTCTACTGGGAC 60  
QY 61 GTGCTTCAGAGGATATGCTGGGACACATACGCGAGAGATACCGAGTGGTACGAT 120  
DB 61 GTGCTTCAGAGGATATGCTGGGACACATACGCGAGAGATACCGAGTGGTACGAT 120  
QY 121 GCCGGAATCTCCGCAATATGATTCCTCCCGGCGAGCAAGGCGATGGCGGCTTATTCG 180  
DB 121 GCCGGAATCTCCGCAATATGATTCCTCCCGGCGAGCGGGTATGAGCGGCTTATTCG 180  
QY 181 ATGGGCTACAGCCCTTACGACTTCTTTGACCTGCTGGGTAGTACGACAGAGGAAAGCT 240  
DB 181 ATGGGCTACAGCCCTTACGACTTCTTTGACCTGCTGGGTAGTACGACAGAGGAAAGCT 240  
QY 241 GAGACGCGCTTTCGCTCCAGCAGGAGCTCTGCAACATGATAAACAACCGCCACGCTTAT 300  
DB 241 GAGACGCGCTTTCGCTCCAGCAGGAGCTCTGCAACATGATAAACAACCGCCACGCTTAT 300  
QY 301 GGCAATGAAGTATATAGCGGATATAGTCAACACCGCGCGCGGCTGACCTGGAGTGG 360  
DB 301 GGCAATGAAGTATATAGCGGATATAGTCAACACCGCGCGCGGCTGACCTGGAGTGG 360  
QY 361 AACCCCTTCTGTAACGACTATACCTGGACCGACTTCTCAAAGGTCTGCGGTAAATAC 420  
DB 361 AACCCCTTCTGTAACGACTATACCTGGACCGACTTCTCAAAGGTCTGCGGTAAATAC 420  
QY 421 ACGGCCAACTACCTCGACTTCACCGAGACGAGCTTCATCGGGCGATTCGCGAAACATTT 480  
DB 421 ACGGCCAACTACCTCGACTTCACCGAGACGAGCTTCATCGGGCGATTCGCGAAACATTT 480  
QY 481 GGAGGCTATCCGACATATGCGGACCAAGAGCTGGGACCAAGTATGCTTGGGCGGACG 540  
DB 481 GGAGGCTATCCGACATATGCGGACCAAGAGCTGGGACCAAGTATGCTTGGGCGGACG 540  
QY 541 CAGGAGCTACCGGCGATATCTCAGAGCATTCGCGATCGCTGCGCTTCGACTAC 600  
DB 541 CAGGAGCTACCGGCGATATCTCAGAGCATTCGCGATCGCTGCGCTTCGACTAC 600  
QY 601 GTCAAGGCTATGCTCCCTGGGTCTGCTCAAGGAGTGGGCTGAACTGGTGGGAGGCTGGGG 660  
DB 601 GTCAAGGCTATGCTCCCTGGGTCTGCTCAAGGAGTGGGCTGAACTGGTGGGAGGCTGGGG 660  
QY 661 GTTGAGAGTACTGGGACCAACAGTCCGCTGAGCGGCTGCTCAACTGGGCGGCTTCGAGCGGT 720

661 GTTGGGAGTACTGGGACCCCAACGTTGATGCCCTCTCCCTGGGCTTACTCGAGCGC 720  
721 GCCAAGGTCTTTGACTTGGCTCTACTACAGATGATGAGGCTTTGACAAACAAAC 780  
721 GCCAAGGTCTTGCATCTCCGCTCTACTACAGATGATGAGGCTTTGACAAACAAAC 780  
781 ATTCCAGCGCTGCTCTCTCCCTCTCAGAACGGCCAGACTGTGCTCTCCCGGACCGGTC 840  
781 ATTCCAGCGCTGCTCTCTCCCTCTCAGAACGGCCAGACTGTGCTCTCCCGGACCGGTC 840  
841 AAGGCGGTAACTTTGTAGCCAAACCAACGATACCGATATTAATCTGGAACAAAGTATCCAGCC 900  
841 AAGGCGGTAACTTTGTAGCCAAACCAACGATACCGATATTAATCTGGAACAAAGTATCCAGCC 900  
901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATTTCTACCGGACTACGAGGAG 960  
901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATTTCTACCGGACTACGAGGAG 960  
961 TGGCTCAACAAAGATAAGCTCAAGAACCTTCATCTGGATACATGAACTCTGCGGAGGGA 1020  
961 TGGCTCAACAAAGATAAGCTCAAGAACCTTCATCTGGATACATGAACTCTGCGGAGGGA 1020  
1021 AGCACCAGCATAGTCTACTACGATACGATGAACCTCTCTGCTGAGGAGCGGCTACCGG 1080  
1021 AGCACCAGCATAGTCTACTACGATACGATGAACCTCTCTGCTGAGGAGCGGCTACCGG 1080  
1081 GACAGCGGGCTTATACCTACATCAACCTAGGCTCGAGCAAGCGCGGAGGTGGGTT 1140  
1081 GACAGCGGGCTTATACCTACATCAACCTAGGCTCGAGCAAGCGCGGAGGTGGGTT 1140  
1141 TATGTCCGAAGTCTCGCGGCGGCTGATCCAGAGTATAGTGTAACTCGGAGGCTGG 1200  
1141 TATGTCCGAAGTCTCGCGGCGGCTGATCCAGAGTATAGTGTAACTCGGAGGCTGG 1200  
1201 GTAGCAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTTCCAGCTTACGACCT 1260  
1201 GTAGCAAGTACGTCTACTCAAGCGGCTGGGTCTATCTCGAAGCTTCCAGCTTACGACCT 1260  
1261 GCCAAGGCGAGTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGGTGA 1311  
1261 GCCAAGGCGAGTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGGTGA 1311

RESULT 12  
US-10-081-872-29  
; Sequence 29, Application US/10081872  
; Publication No. US20030125534A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Prev, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Slupsko, Malgorzata  
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/081,872  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 1311  
; TYPE: DNA

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-29  
Query Match 90.2%; Score 1183; DB 15; Length 1311;  
Best Local Similarity 93.9%; Pred. No. 0;  
Matches 1231; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
Qy 1 ATGGCCAGTACTCCGAGCTGGAAAAGGGCGGGGTCTAATATGCAAGGCTTCTACTGGGAC 60  
Db 1 ATGGCCAGTACTCCGAGCTCGAAGAGGGCGGGGTCTAATATGCAAGGCTTCTACTGGGAC 60  
Qy 61 GTGGCTTTCAGAGGAATATGTGTGGACACAATATCGGCAGAGATACCGGAGTGTGTACGAT 120  
Db 61 GTGGCTTTCAGAGGAATATGTGTGGACACAATATCGGCAGAGATACCGGAGTGTGTACGAT 120  
Qy 121 GCCGGAATCTCCGCAATATGATTTCCCGCGGAGCAAGGGCATGCGCGGCTTATTCG 180  
Db 121 GCCGGAATTTCCGCAATATGATTTCCCGCGGAGCAAGGGCATGCGCGGCTTATTCG 180  
Qy 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAAACGGTA 240  
Qy 241 GAGAGCGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAAACACCGCCACGCTAT 300  
Db 241 GAGAGCGGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATGATAAAACACCGCCACGCTAT 300  
Qy 301 GGCATGAGATATAGCCGATATAGTATCAACACCGCGCGGCGGTGACCTTGGAGTGG 360  
Db 301 GGCATGAGATATAGCCGATATAGTATCAACACCGCGCGGCGGTGACCTTGGAGTGG 360  
Qy 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTTCGCTCGGTAATAT 420  
Db 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTTCGCTCGGTAATAT 420  
Qy 421 ACGGCAACTACTCTGCACTTCCACCGCAACGAGCTCCATGCGGCGGATTCGGGAAACATTT 480  
Db 421 ACTGCCAACTACTCTGCACTTCCACCGCAACGAGCTCCATGCGGCGGATTCGGGAAACATTT 480  
Qy 481 GGAGGCTATCCGACATATGCGACGACGAGCTGGGACGAGTACTGCTTGGGCGGACG 540  
Db 481 GGAGGCTATCCGACATATGCGACGACGAGCTGGGACGAGTACTGCTTGGGCGGACG 540  
Qy 541 CAGGAGGCTACGCGGATATCTCAGGAGCATCGGATCGATCGCTTGGGCTTTCGACTAC 600  
Db 541 CAGGAGGCTACGCGGATATCTCAGGAGCATCGGATCGATCGCTTGGGCTTTCGACTAC 600  
Qy 601 GTCAAGGCTATGCTCCCTGGGTCTGTCAGGAGTGTGCTGAACTGTGTGGGAGGCTGGGG 660  
Db 601 GTCAAGGCTATGCTCCCTGGGTCTGTCAGGAGTGTGCTGAACTGTGTGGGAGGCTGGGG 660  
Qy 661 GTTGGAGAGTACTGGGACACCAACGCTCGACGCTTCTCAAAGTGTGGTGTGGGAGGCTGGGG 720  
Db 661 GTTGGAGAGTACTGGGACACCAACGCTCGACGCTTCTCAAAGTGTGGTGTGGGAGGCTGGGG 720  
Qy 721 GCCAAGGTCTTTGACTTTCGCTTCTACTAAGAGTGTGAGGCTTTCGACCAACAAAC 780  
Db 721 GCCAAGGTCTTTGACTTTCGCTTCTACTAAGAGTGTGAGGCTTTCGACCAACAAAC 780  
Qy 781 ATTCCAGCGCTGCTCTGCTGCTTCCGCTTCCGAGGCGGAGCTGTGCTCTCCCGGACCGGTC 840  
Db 781 ATTCCAGCGCTGCTCTGCTGCTTCCGCTTCCGAGGCGGAGCTGTGCTCTCCCGGACCGGTC 840  
Qy 841 AAGGCGGTAACTTTGTAGCCAAACCAACGATATTAATCTGGAACAAAGTATCCAGCC 900  
Db 841 AAGGCGGTAACTTTGTAGCCAAACCAACGATATTAATCTGGAACAAAGTATCCAGCC 900  
Qy 901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATTTCTACCGGACTACGAGGAG 960  
Db 901 TAGCGCTTTCATCTCTACCTACGAGGCGCCAGCGCAATATTTCTACCGGACTACGAGGAG 960

```
QY 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACTCGCCCGAGGA 1020
Db 961 TGGCTCAACAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACTCGCCCGAGGA 1020

QY 1021 AGCACCGACATAGTCTACTACGATACGATGAACTCATCTCGTCAAGAAAGGCTACGGG 1080
Db 1021 AGCACCGACATAGTCTACTACGATACGATGAACTCATCTCGTCAAGAAAGGCTACGGG 1080

QY 1081 GACAAAGCGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGGAAGTGGTT 1140
Db 1081 GACAAAGCGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGGAAGTGGTT 1140

QY 1141 TATGTCCGGAAGTTCGCGGGCGGCTGCATCCAGAGTATCTGGTAAACCTCGAGGCTGG 1200
Db 1141 TATGTCCGGAAGTTCGCGGGCGGCTGCATCCAGAGTATCTGGTAAACCTCGAGGCTGG 1200

QY 1201 GTAGACAAGTACGCTCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260
Db 1201 GTAGACAAGTACGCTCTACTCAAGCGGCTGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260

QY 1261 GCCAAAGCGGAGTATGGCTACTCCGTGTGAGCTACTCGCGGGTGGGCTGA 1311
Db 1261 GCCAAAGCGGAGTATGGCTACTCCGTGTGAGCTACTCGCGGGTGGGCTGA 1311

RESULT 13
US-10-385-305-29
; Sequence 29 Application US/10385305
; Publication No. US20040018607A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kerovuo, James S.
; APPLICANT: Slupaska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: synthetically generated oligonucleotide
US-10-385-305-29

Query Match 90.2%; Score 1183; DB 16; Length 1311;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 ATGCCCAAGTACTCCGAGCTGGAAAGCGGGGTCTATATGCAAGGGTCTTACTGGGAC 60
Db 1 ATGCCCAAGTACTCCGAGCTGGAAAGCGGGGTCTATATGCAAGGGTCTTACTGGGAC 60

QY 61 GTGCTTCAGGAGGAATATGTGGGACACAAATACGGCAGAGATACCGAGTGTGACAT 120
Db 61 GTGCTTCAGGAGGAATATGTGGGACACAAATACCGAGTGTGACAT 120
```

```
QY 121 GCCGGAATCTCCCAATATATGATTTCCCGCGGAGCAAGGGATGGCGGCGCTATTTCG 180
Db 121 GCCGGAATCTCCCGCATATGGAATTCGCCCGGAGCAAGGGATGGCGGCGCTATTTCG 180

QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAAAGGTA 240
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGAAAGGTA 240

QY 241 GAGACGCGCTTTGGCTCCAGCAGGAGCTGCTGAACATGATAAACCACCGCCACGCTAT 300
Db 241 GAGACGCGCTTTGGCTCCAGCAGGAGCTGCTGAACATGATAAACCACCGCCACGCTAT 300

QY 301 GGATAGAGTATAGCCGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360
Db 301 GGATAGAGTATAGCCGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360

QY 361 AACCCCTTCGTGAACGACTATACCTGACCGACTTCTCAAAGGTGCGTGGGTAAATAC 420
Db 361 AACCCCTTCGTGGGACTACACTGACCGACTTCTCAAAGGTGCGTGGGCAATAT 420

QY 421 ACGCCCAACTACCTCGACTTCCACCGAACGAGCTCCATCGCGCGGATTCGGGAACATTT 480
Db 421 ACTGCCAACTACCTCGACTTCCACCGAACGAGCTCCATCGCGCGGATTCGGGAACATTT 480

QY 481 GGAGGCTATCCGACATATGCCACGAGCTGGGACCCAGTACTGGCTCTCGGCGAGC 540
Db 481 GGAGGCTATCCGACATATGCCACGAGCTGGGACCCAGTACTGGCTCTCGGCGAGC 540

QY 541 CAGGAGAGTACGCGGATATCTCAGGAGCATCGGCAATCGCTGGCGCTTCGACTAC 600
Db 541 CAGGAGAGTACGCGGATATCTCAGGAGCATCGGCAATCGCTGGCGCTTCGACTAC 600

QY 601 GTCAAGGCTATGCTCCCTGGGTGCTCAAGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660
Db 601 GTCAAGGCTATGCTCCCTGGGTGCTCAAGAGCTGGCTGAACTGGTGGGAGGCTGGGCG 660

QY 661 GTTGAGAGTACTGGGACACCAACGCTGAGCTGTTCTCAACTGGGATACTCGAGCGGT 720
Db 661 GTTGAGAGTACTGGGACACCAACGCTGAGCTGTTCTCAACTGGGATACTCGAGCGGT 720

QY 721 GCCAAGGCTTTTGAATTCGCGCTTACTACAAGTGAATGAGGCTTTGACAAACAAAC 780
Db 721 GCCAAGGCTTTTGAATTCGCGCTTACTACAAGTGAATGAGGCTTTGACAAACAAAC 780

QY 781 ATTCAGGCTCGTCTCTGCCCTTCAGAAAGCGGACAGCTGTTGTCTCCCGGACCCGTT 840
Db 781 ATTCAGGCTCGTCTCTGCCCTTCAGAAAGCGGACAGCTGTTGTCTCCCGGACCCGTT 840

QY 841 AAGCGCGTAACTTTGTAGAAAACACGACACCGATATATCTGGAACAAGTATCCAGCC 900
Db 841 AAGCGCGTAACTTTGTAGAAAACACGACACCGATATATCTGGAACAAGTATCCAGCC 900

QY 901 TAGCGTTTCATCTCCTCAGGAGGCGGAGCAATATTTCTACCGGACTACGAGGAG 960
Db 901 TAGCGTTTCATCTCCTCAGGAGGCGGAGCAATATTTCTACCGGACTACGAGGAG 960

QY 961 TGGCTCAAAGGATAAGCTCAAGAACTCATCTGGATATATGAGAACCTCGCGGAGGA 1020
Db 961 TGGCTCAAAGGATAAGCTCAAGAACTCATCTGGATATATGAGAACCTCGCGGAGGA 1020

QY 1021 AGCACCGACATAGTCTACTACGATACGATGAACTCATCTTCGTCAGAAAGGCTACGGG 1080
Db 1021 AGCACCGACATAGTCTACTACGATACGATGAACTCATCTTCGTCAGAAAGGCTACGGG 1080

QY 1081 GACAAAGCGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGGAAGTGGTT 1140
Db 1081 GACAAAGCGGGCTTATAACCTACATCAACCTAGGCTCGAGCAAGCGGGAAGTGGTT 1140

QY 1141 TATGTCCGGAAGTTCGCGGGCGGCTGCATCCAGAGTATCTGGTAAACCTCGAGGCTGG 1200
Db 1141 TATGTCCGGAAGTTCGCGGGCGGCTGCATCCAGAGTATCTGGTAAACCTCGAGGCTGG 1200

QY 1201 GTAGACAAGTATGGCTACTCCGTGTGAGCTACTCGCGGGTGGGCTGA 1260
Db 1201 GTAGACAAGTATGGCTACTCCGTGTGAGCTACTCGCGGGTGGGCTGA 1260
```

Db 1201 GTAGCAAGTACCTCGATCCAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGTACTCGGGGTGGGCTGA 1311  
Db 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGTACTCGGGGTGGGCTGA 1311

RESULT 14

US-10-081-872-13  
; Sequence 13, Application US/10081872  
; Publication No. US2003012534A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.  
; APPLICANT: Kerovuo, Janne S.  
; APPLICANT: Slupeksa, Malgorzata  
; TITLE OF INVENTION: ENZYMS HAVING ALPHA AMYLASE ACTIVITY  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; FILE REFERENCE: 09010-108001  
; CURRENT APPLICATION NUMBER: US/10/081,872  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,495  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/270,496  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 1311  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-10-081-872-13

Query Match 89.4%; Score 1171.8; DB 15; Length 1311;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1224; Conservative 0; Mismatches 87; Indels 0; Gaps 0;  
QY 1 ATGGCCAAAGTACTCGAGCTGGAAGGGCGGGGTCAATCGCAGGCGTCTACTGGGAC 60  
Db 1 ATGGCCAAAGTACTCGAGCTGGAAGGGCGGGGTCAATCGCAGGCGTCTACTGGGAC 60  
QY 61 GTGCTTCAGGAGGAATATGGTGGGACACAATACGGCAGAGATACCGGAGTGTACGAT 120  
Db 61 GTGCTTCAGGAGGAATATGGTGGGACACAATACGGCAGAGATACCGGAGTGTACGAT 120  
QY 121 GCCGGAATCTCGGCAATATGATATCCCGCGGAGCAAGGCGGCGGCGCTATTTCG 180  
Db 121 GCCGGAATCTCGGCAATATGATATCCCGCGGAGCAAGGCGGCGGCGCTATTTCG 180  
QY 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAAGAGGGAACGGTA 240  
Db 181 ATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTATGACCAAGAGGGAACGGTA 240  
QY 241 GAGACGGCTTTGGTCCAGCAGAGAGCTGTGAACATGATATAACACCGCCCGACGCTAT 300  
Db 241 GAGACGGCTTTGGTCCAGCAGAGAGCTGTGAACATGATATAACACCGCCCGACGCTAT 300  
QY 301 GGCATGAAGGTAATAGCGGATATAGTATCAACACCGCGCGGCGGTGACCTGGAGTGG 360  
Db 301 GGCATGAAGGTCATAGCGGATATAGTATCAACACCGCGCGGCGGTGACCTGGAGTGG 360  
QY 361 AACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGTTCGGGTAAATAC 420  
Db 361 AACCCCTTCGTTCGGGACTACACCTGGACCGACTTCTCAAAGGTTCGGGTAAATAT 420

QY 421 ACGCCCAACTACCTCGATCCAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 480  
Db 421 ACTGCCAACTACCTCGATCCAGCGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 480  
QY 481 GGAGGCTATCCCGACATATGCGCAGCAAGAGCTGGGACCAAGTACTGGGCGGAGC 540  
Db 481 GGAGGCTATCCCGACATATGCGCAGCAAGAGCTGGGACCAAGTACTGGGCGGAGC 540  
QY 541 CAGGAGAGCTACCGGCGATATCTCAGGAGCATCGGATCGATCGCTGGGCGGCTGACTAC 600  
Db 541 GATGAGAGCTACCGGCGATATCTCAGGAGCATCGGATCGATCGCTGGGCGGCTGACTAC 600  
QY 601 GTCAAGGGCTATGCTCCCTGGGTCTGCTCAAGGACTTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
Db 601 GTCAAGGGCTACGGAGCGTGGGTCTGCTCAAGGACTTGGCTGAACCTGGTGGGAGGCTGGGCG 660  
QY 661 GTTGGAGTACTGGGACACCAAGCTCGAGCTGTTCTCAACTGGGCACTACTCGAGGGT 720  
Db 661 GTTGGAGTACTGGGACACCAAGCTCGAGCTGTTCTCAACTGGGCACTACTCGAGGGT 720  
QY 721 GCCAAGGTCTTTGACTTGGGCTCTACTCAAGATGGATGAGGCTTTGACAAACAAAAC 780  
Db 721 GCAAAAGTCTTCGACTTCCGCTCTACTCAAGATGGATGAGGCTTTGACAAACAAAAC 780  
QY 781 ATCCAGCGCTCGTCTCTGCTTCAAGCGGCGAGCTGTTCTCTCCGCGGACCGGCTTC 840  
Db 781 ATCCAGCGCTCGTCTCTGCTTCAAGCGGCGAGCTGTTCTCTCCGCGGACCGGCTTC 840  
QY 841 AAGGCGGTAACCTTTGTAGCAACACGACGATATATCTGGAACAGTATCCAGGCT 900  
Db 841 AAGGCGGTAACCTTTGTAGCAACACGACGATATATCTGGAACAGTATCCAGGCT 900  
QY 901 TAGCGGTTATCTCTACCTACGAGGCGCGGCAATATTTCTACCGGCTACTACGAGGAG 960  
Db 901 TAGCGGTTATCTCTACCTACGAGGCGCGGCAATATTTCTACCGGCTACTACGAGGAG 960  
QY 961 TGGCTCAACAGGATAGCTCAGAACCTCATCTGATATGATGAGAACCTTCGCGCGGAGGA 1020  
Db 961 TGGCTCAACAGGATAGCTCAGAACCTCATCTGATATGATGAGAACCTTCGCGCGGAGGA 1020  
QY 1021 AGCACGACATAGTCTACTACGATACGATGAACCTCATCTTCTGTCAGGAAAGGCTACGGG 1080  
Db 1021 AGCACTGACATAGTCTACTACGATACGATGAACCTCATCTTCTGTCAGGAAAGGCTACGGG 1080  
QY 1081 GACAAGCGGGGCTTATACCTACATCAACCTAGGCTCGAGCAAGCGCGGAGGTTGGTT 1140  
Db 1081 GACAAGCGGGGCTTATACCTACATCAACCTAGGCTCGAGCAAGCGCGGAGGTTGGTT 1140  
QY 1141 TATGTCCGAAGTTCCGCGGCGGTCATCCAGAGTATCTGTTAACTCCGAGGCTCG 1200  
Db 1141 TATGTCCGAAGTTCCGCGGCGGTCATCCAGAGTATCTGTTAACTCCGAGGCTCG 1200  
QY 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
Db 1201 GTAGCAAGTACGCTTACTCAAGCGGCTGGGTCTATCTCGAAGCTCCAGCTTACGACCT 1260  
QY 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGGCTGA 1311  
Db 1261 GCCAAGCGGAGTATGGCTACTCCGTGTGGAGCTACTCGGGGTGGGCTGA 1311

RESULT 15

US-10-385-305-13  
; Sequence 13, Application US/10385305  
; Publication No. US20040018607A1  
; GENERAL INFORMATION:  
; APPLICANT: Callen, Walter  
; APPLICANT: Richardson, Toby  
; APPLICANT: Frey, Gerhard  
; APPLICANT: Short, Jay M.  
; APPLICANT: Mathur, Eric J.  
; APPLICANT: Gray, Kevin A.



APPLICANT: Kerovuo, Janne S.  
APPLICANT: Slupska, Malgorzata  
TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
TITLE OF INVENTION: AND METHODS OF USE THEREOF  
FILE REFERENCE: 09010-108001  
CURRENT APPLICATION NUMBER: US/10/385,305  
CURRENT FILING DATE: 2003-03-06  
PRIOR APPLICATION NUMBER: US/10/081,872  
PRIOR FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: US 60/270,495  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: US 60/270,496  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: US 60/291,122  
PRIOR FILING DATE: 2001-05-14  
NUMBER OF SEQ ID NOS: 321  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 1311  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetically generated oligonucleotide

Query Match 89.4%; Score 1171.8; DB 16; Length 1311;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1224; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```
Qy 1 ATGCCAAGTACCTCCAGAGTGGAAAAGGGGGGGTCAATATCGAGCGTTCTACTGGGAC 60
Db 1 ATGCCAAGTACCTCCAGAGTGGAAAAGGGGGGGTCAATATCGAGCGTTCTACTGGGAC 60
Qy 61 GTCCCTTCAGAGGAAATATGCTGGGACACAATACGCGAGAGATACCGGAGTGGTACGAT 120
Db 61 GTCCCTTCAGAGGAAATATGCTGGGACACAATACGCGAGAGATACCGGAGTGGTACGAT 120
Qy 121 GCGGGAATCTCCGAATATAGTATCCCGGCGAGCAAGGCGATGGGGCGGCTATTTCG 180
Db 121 GCGGGAATCTCCGAATATAGTATCCCGGCGAGCAAGGCGATGGGGCGGCTATTTCG 180
Qy 181 ATGGGCTACACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGCAACGGTA 240
Db 181 ATGGGCTACACCCCTACGACTTCTTTGACCTCGGTGAGTACGACAGAGGCAACGGTA 240
Qy 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATATGATAAACACCGCCACGCTAT 300
Db 241 GAGACGCGCTTTGGCTCCAAAGCAGGAGCTCGTGAACATATGATAAACACCGCCACGCTAT 300
Qy 301 GGCATGAGGTATAGCCGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360
Db 301 GGCATGAGGTATAGCCGATATAGTATCAACACCGCGCGCGGTGACCTGGAGTGG 360
Qy 361 AACCCCTTCGTGACGACTATACCTGACCGGACTTCTCAAAGGTGCGCTGGGTAAATAC 420
Db 361 AACCCCTTCGTGACGACTATACCTGACCGGACTTCTCAAAGGTGCGCTGGGTAAATAC 420
Qy 421 ACGGCCAACTACCTCGACTTCCACCCGAAAGAGTCCATCGCGGCGGATTCGGGAACATTT 480
Db 421 ACTGCCAACTACCTCGACTTCCACCCGAAAGAGTCCATCGCGGCGGATTCGGGAACATTT 480
Qy 481 GGRGGCTATCCGACATATCCAGCAAGAGCTGGGACCGAGTACTGGCTCTGGGCGAGC 540
Db 481 GGRGGCTATCCGACATATCCAGCAAGAGCTGGGACCGAGTACTGGCTCTGGGCGAGC 540
Qy 541 CAGGAGAGCTACCGGCGATATCTCAGGAGCATCGGCATCGCATCGCTGGCGCTTCGACTAC 600
Db 541 GATGAGAGCTACCGGCGCTACCTAAGGAGCATCGGCGCTTCATGCTGGCGCTTCGACTAC 600
Qy 601 GTCAGGGCTATGCTCCCTGGGTGCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGCG 660
Db 601 GTCAGGGCTATGCTCCCTGGGTGCTCAAGGAGCTGGCTGAACTGGTGGGAGGCTGGCG 660
```

Search completed: June 29, 2004, 13:31:50  
Job time : 611 secs

```
Qy 661 GTTGGAGAGTACTGGGACACCAACGCTGAGCGTGTCTCAACTGGGCATATCTGAGCGGT 720
Db 661 GTTGGAGAGTACTGGGACACCAACGCTGAGCGTGTCTCAACTGGGCCTACTCTGAGCGGT 720
Qy 721 GCCAAGGTCTTTGACTTGGCCCTCTACTACAGATGGATGAGCCCTTTGACACCAAAAC 780
Db 721 GCCAAGGTCTTTGACTTGGCCCTCTACTACAGATGGATGAGCCCTTTGACACCAAAAC 780
Qy 781 ATTCCAGCGCTGCTCTGCTGCCCTTCAGAACGGCCAGACTGTCTCTCCCGCGACCCGTTTC 840
Db 781 ATTCCAGCGCTGCTCTCTGCGCTTCAGAACGGCCAGACTGTCTCTCCCGCGACCCGTTTC 840
Qy 841 AAGCCGTAACCTTTGTAGCAAAACCCAGCACCGATATATCTGGAACAAAGTATCCAGCC 900
Db 841 AAGCCGTAACCTTTGTAGCAAAACCCAGCACCGATATATCTGGAACAAAGTATCCAGCC 900
Qy 901 TAGCGGTTTCATCTCTACCTACGAGGGCCAGCCGACAAATATTTCTACCGCGACTACGAGGAG 960
Db 901 TAGCGGTTTCATCTCTACCTACGAGGGCCAGCCGACAAATATTTCTACCGCGACTACGAGGAG 960
Qy 961 TGGCTCAACAAAGGATAGCTCAAGAAACCTCATCTGGATACATGAGAACTCTCGCGGAGGA 1020
Db 961 TGGCTCAACAAAGGATAGCTCAAGAAACCTCATCTGGATACATGAGAACTCTCGCGGAGGA 1020
Qy 1021 AGCACCGACATAGTCTTACTAGTACGATGAACTCATCTCTGTCAGGAAACGCTACCGG 1080
Db 1021 AGCACCGACATAGTCTTACTAGTACGATGAACTCATCTCTGTCAGGAAACGCTACCGG 1080
Qy 1081 GACAAAGCGGGGCTTATAAACCCTACATCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGTT 1140
Db 1081 GACAAAGCGGGGCTTATAAACCCTACATCAACCTAGGCTCGAGCAAGGCCGGAAGGTGGTT 1140
Qy 1141 TATGTGCCGAAGTTTCGGGGCGGTGTCATCCACGAGTATCTGTTAACTCTCGGAGGCTGG 1200
Db 1141 TATGTGCCGAAGTTTCGGGGCGGTGTCATCCACGAGTATCTGTTAACTCTCGGAGGCTGG 1200
Qy 1201 GTAGACAAGTACGTCTACTCAAAGCGGTGGGTCTATCTCGAAGCTCCAGCTTACGACCCCT 1260
Db 1201 GTAGACAAGTACGTCTACTCAAAGCGGTGGGTCTATCTCGAAGCTCCAGCTTACGACCCCT 1260
Qy 1261 GCCAAGCGGCAAGTATGCTACTCCGTGTGAGCTACTGCGGGGTGGGCTGA 1311
Db 1261 GCCAAGCGGCAAGTATGCTACTCCGTGTGAGCTACTGCGGGGTGGGCTGA 1311
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:29:46 ; Search time 59 Seconds  
(without alignments)  
2087.978 Million cell updates/sec

Title: US-10-081-739A-2

Perfect score: 2459

Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQGVSVMSYCGVG 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A: Genesep25Jan04: \*  
1: Genesep1980s: \*  
2: Genesep1990s: \*  
3: Genesep2000s: \*  
4: Genesep2001s: \*  
5: Genesep2002s: \*  
6: Genesep2003as: \*  
7: Genesep2003bs: \*  
8: Genesep2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2459	100.0	436	6	ABG70531 Alpha amy
2	2459	100.0	436	6	ABP70154 Amino aci
3	2424	98.6	436	6	ABU03074 Alpha amy
4	2406	97.8	436	6	ABU03042 Alpha amy
5	2378	96.7	436	6	ABU03068 Alpha amy
6	2377	96.7	436	6	ABU03057 Alpha amy
7	2371	96.4	436	6	ABU03053 Alpha amy
8	2368	96.3	436	6	ABU03041 Alpha amy
9	2360	96.0	436	6	ABU03058 Alpha amy
10	2358	95.9	436	6	ABU03046 Alpha amy
11	2358	95.9	436	6	ABU03069 Alpha amy
12	2356	95.8	436	6	ABU03054 Alpha amy
13	2356	95.8	436	6	ABU03048 Alpha amy
14	2352	95.6	436	6	ABU03066 Alpha amy
15	2350	95.6	436	6	ABU03043 Alpha amy
16	2350	95.6	436	6	ABU03052 Alpha amy
17	2349	95.5	436	6	ABP96602 797GL3 al
18	2349	95.5	436	6	ABU03059 Alpha amy
19	2349	95.5	436	6	ABU03061 Alpha amy
20	2349	95.5	518	6	ABP96609 Plasmid p
21	2349	95.5	741	6	ABP96606 Maize 797
22	2349	95.5	820	6	ABP96610 Plasmid p
23	2347	95.4	436	6	ABU03045 Alpha amy
24	2344	95.3	454	6	ABP96607 Plasmid p
25	2344	95.3	460	6	ABP96622 797GL3 fu

26	2344	95.3	460	6	ABP96608 Plasmid p
27	2344	95.3	460	6	ABP96624 797GL3 fu
28	2342	95.2	436	6	ABU03051 Alpha amy
29	2338	95.1	436	6	ABU03044 Alpha amy
30	2338	95.1	436	6	ABU03050 Alpha amy
31	2334	94.9	436	6	ABU03064 Alpha amy
32	2334	94.9	436	6	ABU03071 Alpha amy
33	2332	94.8	436	6	ABU03040 Alpha amy
34	2331	94.8	436	6	ABU03049 Alpha amy
35	2331	94.8	436	6	ABU03056 Alpha amy
36	2327	94.6	436	6	ABU03070 Alpha amy
37	2321	94.4	436	6	ABU03055 Alpha amy
38	2314	94.1	432	6	ABG70536 Alpha amy
39	2314	94.1	432	6	ABU03076 Alpha amy
40	2311	94.0	432	6	ABU03077 Alpha amy
41	2298	93.5	435	6	ABU03065 Alpha amy
42	2298	93.5	461	6	ABG70537 Alpha amy
43	2298	93.5	461	6	ABU03078 Alpha amy
44	2290	93.1	436	6	ABU03060 Alpha amy
45	2283	92.8	436	6	ABU03062 Alpha amy

## ALIGNMENTS

RESULT 1  
ABG70531  
ID ABG70531 standard; protein; 436 AA.

AC ABG70531;

DT 20-JAN-2003 (first entry)

DE Alpha amylase polypeptide #1.

KW Alpha amylase; enzyme; starch liquefaction; starch linkage hydrolysis;  
KW starch breakdown catalysis; high-maltose syrup; high-glucose syrup; rice;  
KW germinated rice; corn; barley; wheat; legume; sweet potato; anti-staling;  
KW textile desizing; lignocellulosic fibre treatment; enzymatic de-linking;  
KW recycled paper pulp; alpha-glucosidic linkage; dough; crumb firming;  
KW crumb elasticity; palatability; improved flavour.

OS Unidentified.

XX WO200268597-A2.

XX 06-SEP-2002.

XX 21-FEB-2002; 2002WO-US005538.

XX 21-FEB-2001; 2001US-0270495P.

XX 21-FEB-2001; 2001US-0270496P.

XX 14-MAY-2001; 2001US-0291122P.

(DIVE-) DIVERSA CORP.

XX Callen W, Richardson T, Frey G, Miller C, Kazaoka M, Mathur EJ;

XX Short JM;

XX WPI; 2003-018657/01.

XX N-PSDB; ABS55928.

PT Novel alpha amylase polypeptide, useful e.g. for starch liquefaction,

PT textile desizing, treating lignocellulosic fibers and for preparing dough

PT or baked product, has increased activity and stability.

XX Claim 30; Fig 7; 147pp; English.

XX The invention relates to a purified alpha amylase polypeptide and the

XX polynucleotide encoding it. The polypeptide is useful for liquefying a

XX starch containing composition, hydrolysing a starch linkage, catalysing

XX the breakdown of a starch and producing high-maltose or high-glucose

XX syrup or mixed syrup, where the starch is from rice, germinated rice,

corn, barley, wheat, legumes or sweet potato. The polypeptide is also useful for washing an object, textile desizing, treating lignocellulosic fibres for improving fibre properties, enzymatic de-inking of recycled paper pulp and modifying small molecules by biocatalytic reaction. The polypeptide is further useful for increasing the flow of production fluids from a subterranean formation by removing a viscous, starch-containing, damaging fluid formed during production operations and found within the subterranean formation which surrounds a completed well bore, by pumping an enzyme treatment comprising the polypeptide and an aqueous fluid to a desired location within the well-bore which is effective to attack the alpha-glucosidic linkages in the starch-containing fluid. The alpha amylases have increased activity and stability at increased pH and temperature and can be used in preparing dough to provide an improved anti-staling effect as measured by less crumb firming, retained crumb elasticity, improved slice-ability, improved palatability or improved flavour. This sequence represents an alpha amylase of the invention

XX SQ Sequence 436 AA;

Query Match 100.0%; Score 2459; DB 6; Length 436;  
Best Local Similarity 100.0%; Pred. No. 3.3e-203;  
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPPASKMGGAYS 60  
DB 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPPASKMGGAYS 60  
QY 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
DB 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
DB 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
QY 181 QESYAALYRSIGIDAWRFDYKGYAPVVKDNLNMGWGAUGEYWDITNDAVLNWAYSSG 240  
DB 181 QESYAALYRSIGIDAWRFDYKGYAPVVKDNLNMGWGAUGEYWDITNDAVLNWAYSSG 240  
QY 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
DB 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
QY 301 YAFILTYEGQPTIFPYRDEEELNKKLNIWIHENLAGGSTDIVYDNDDELIFVRNGYG 360  
DB 301 YAFILTYEGQPTIFPYRDEEELNKKLNIWIHENLAGGSTDIVYDNDDELIFVRNGYG 360  
QY 361 DKPGLITVINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYYVSSGWVYLEAPAYDP 420  
DB 361 DKPGLITVINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYYVSSGWVYLEAPAYDP 420  
QY 421 ANGOYGYSVMSYCGVG 436  
DB 421 ANGOYGYSVMSYCGVG 436

RESULT 2  
ABP70154  
ID ABP70154 standard; protein; 436 AA.  
XX AC ABP70154;  
XX DT 07-APR-2003 (first entry)  
XX DE Amino acid sequence of an alpha-amylase.  
XX KW Alpha-amylase; enzyme; detergent; baking; beverage; fuel; ethanol.  
XX CS Bacillus sp.  
XX FN WO200292802-A1.  
XX PD 21-NOV-2002.

XX 14-MAY-2002; 2002WO-US015422.  
XX 14-MAY-2001; 2001US-0291122P.  
XX (DIVE-) DIVERSA CORP.  
XX Gerendash J;  
XX WPI; 2003-156728/15.  
XX N-FSDS; ABZ33371.  
XX Purifying enzyme, such as alpha-amylase and cellulase, involves  
PT flocculating fermentation broth containing bacterial cells containing the  
PT enzyme, centrifuging and extracting enzymes by filtration.  
XX Disclosure; Fig 7; 108pp; English.

The present sequence represents an alpha-amylase enzyme, which may be purified using the method of the invention. The specification describes a method for purifying an enzyme, such as alpha-amylase and cellulase, by flocculating fermentation broth containing bacterial cells containing the enzyme, centrifuging and extracting enzymes by filtration. Alternatively, the above method involves subjecting a fermentation broth to a heat-killing procedure, where a resultant heat-killed broth is formed, washing the heat-killed broth with a buffered medium, releasing the enzyme contained in the cells of the fermentation broth and extracting the enzyme by microfiltration. The method is useful for purifying an enzyme derived from a mixed population of organisms or an isolate from the population. The alpha-amylases are useful in corn-wet milling processes, detergents, baking processes, beverages and in oil fields (fuel ethanol)

XX SQ Sequence 436 AA;

Query Match 100.0%; Score 2459; DB 6; Length 436;  
Best Local Similarity 100.0%; Pred. No. 3.3e-203;  
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPPASKMGGAYS 60  
DB 1 MAKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPPASKMGGAYS 60  
QY 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
DB 61 MGYPDPDFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
DB 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWAS 180  
QY 181 QESYAALYRSIGIDAWRFDYKGYAPVVKDNLNMGWGAUGEYWDITNDAVLNWAYSSG 240  
DB 181 QESYAALYRSIGIDAWRFDYKGYAPVVKDNLNMGWGAUGEYWDITNDAVLNWAYSSG 240  
QY 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
DB 241 AKVDFPDLALYKMDAEPFNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
QY 301 YAFILTYEGQPTIFPYRDEEELNKKLNIWIHENLAGGSTDIVYDNDDELIFVRNGYG 360  
DB 301 YAFILTYEGQPTIFPYRDEEELNKKLNIWIHENLAGGSTDIVYDNDDELIFVRNGYG 360  
QY 361 DKPGLITVINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYYVSSGWVYLEAPAYDP 420  
DB 361 DKPGLITVINLGSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYYVSSGWVYLEAPAYDP 420  
QY 421 ANGOYGYSVMSYCGVG 436  
DB 421 ANGOYGYSVMSYCGVG 436

RESULT 3  
ABUC3074

ID ABU03074 standard; protein; 436 AA.  
 XX  
 AC ABU03074;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Alpha amylase polypeptide #35.  
 XX  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 XX starch modification.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200269589-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 21-FEB-2002; 2002WO-US005068.  
 XX  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX  
 PA (DIVE-) DIVERSA CORP.  
 XX  
 PI Callen W, Richardson T, Frey G;  
 XX  
 FI WPI; 2003-018656/01.  
 DR N-PSDB; ABX08445.  
 DR  
 DR  
 XX  
 PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 PS  
 PS Claim 30; Fig 16; 301pp; English.  
 XX  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX  
 SQ Sequence 436 AA;  
 Query Match 98.6%; Score 2424; DB 6; Length 436;  
 Best Local Similarity 98.4%; Pred. No. 3.4e-200;  
 Matches 429; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MAKYLELEGGVIMQAFYWDYVPSGGIWDITRQKIPEDYDAGISAIWTPPASKMGWGAYS 60  
 DB 1 MAKYLELEGGVIMQAFYWDYVPSGGIWDITRQKIPEDYDAGISAIWTPPASKMGWGAYS 60  
 QY 61 MGYPDPDFDLGVDYDQKVTETRGSKOELVNMINTAHAYGMKIADIVINHRAGGDLW 120  
 DB 61 MGYPDPDFDLGVDYDQKVTETRGSKOELVNMINTAHAYGMKIADIVINHRAGGDLW 120

QY 121 NPFVNDYTDWTFDSKVASSGKYTANYLDFHPNELHAGDSGTFGGYPDICHKSDQYWLWAS 180  
 DB 121 NPFVNDYTDWTFDSKVASSGKYTANYLDFHPNELHAGDSGTFGGYPDICHKSDQYWLWAS 180  
 QY 181 QESYAAYLRISIGIDAWRFDYVKGYPVWVKDNLNMGCGWAGVGEYWDNTVDVNLWAYSSG 240  
 DB 181 QESYAAYLRISIGIDAWRFDYVKGYPVWVKDNLNMGCGWAGVGEYWDNTVDVNLWAYSSG 240  
 QY 241 AKVFDFALYYKMDAEAFNKNIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 DB 241 AKVFDFALYYKMDAEAFNKNIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGQPTTFYRDYEEWLNKDKLKNLIWIHNLNLAGGSTDIVVYDNDDELIFVRNGY 360  
 DB 301 YAFILTYEGQPTTFYRDYEEWLNKDKLKNLIWIHNLNLAGGSTDIVVYDNDDELIFVRNGY 360  
 QY 361 DKPLITIYNILGSSKAGRWVYVPKAGACIHEYTNLGGWVDKVVYSSGWVYLEAPAYDP 420  
 DB 361 SKPLITIYNILGSSKAGRWVYVPKAGACIHEYTNLGGWVDKVVYSSGWVYLEAPAYDP 420  
 QY 421 ANGQGYSVWSYCGVG 436  
 DB 421 ANGQGYSVWSYCGVG 436  
 RESULT 4  
 ABU03042  
 ID ABU03042 standard; protein; 436 AA.  
 XX  
 AC ABU03042;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Alpha amylase polypeptide #3.  
 XX  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200268589-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 21-FEB-2002; 2002WO-US005068.  
 XX  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX  
 PA (DIVE-) DIVERSA CORP.  
 XX  
 PI Callen W, Richardson T, Frey G;  
 XX  
 FI WPI; 2003-018656/01.  
 DR N-PSDB; ABX08413.  
 DR  
 DR  
 XX  
 PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 PS  
 PS Claim 30; Fig 16; 301pp; English.  
 XX  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-inking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX  
 SQ Sequence 436 AA;

CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging  
CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of  
CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
CC polypeptides of the invention  
XX  
SQ Sequence 436 AA;

Query Match 97.8%; Score 2406; DB 6; Length 436;  
Best Local Similarity 97.5%; Pred. No. 1.2e-198;  
Matches 425; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAKYSELEGGVIMQAFYNDVPSSGGIWDITIRQKIPEDYDAGISAIWIIPASKMGSGAYS 60  
DB 1 MAKYSELEGGVIMQAFYNDVPSSGGIWDITIRQKIPEDYDAGISAIWIIPASKMGSGAYS 60  
QY 61 MGYPDPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
DB 61 MGYPDPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
QY 121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180  
DB 121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180  
QY 181 QESYAALYRISIGIDAWRFDYKGYAPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
DB 181 QESYAALYRISIGIDAWRFDYKGYAPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
QY 241 AKVPDFALYKMDAEFNNKIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
DB 241 AKVPDFALYKMDAEFNNKIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
QY 301 YAFILTYEGQPTIFRYDYEELNKKLNIWHENLAGSGTDIVYDNDDELIFVRNGYG 360  
DB 301 YAFILTYEGQPTIFRYDYEELNKKLNIWHENLAGSGTDIVYDNDDELIFVRNGYG 360  
QY 421 ANGOYGSVMSYCGVG 436  
DB 421 ANGOYGSVMSYCGVG 436

RESULT 5  
ABU03068  
ID ABU03068 standard; protein; 436 AA.  
XX  
AC ABU03068;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Alpha amylase polypeptide #29.  
XX  
KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
OS Unidentified.  
XX  
FN WO200268589-A2.  
XX

PD 06-SEP-2002.  
XX  
XX 21-FEB-2002; 2002WO-US005068.  
XX  
XX 21-FEB-2001; 2001US-0270495P.  
PR 21-FEB-2001; 2001US-0270496P.  
PR 14-MAY-2001; 2001US-0291122P.  
XX  
XX (DIVE-) DIVERSA CORP.  
XX  
XX Callen W, Richardson T, Frey G;  
XX WPI; 2003-018656/01.  
XX N-PSDB; ABX08439.  
XX  
XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
XX liquefying starch, for textile desizing, for treating lignocellulosic  
XX fibers, and for producing high-maltose or high-glucose syrup.  
XX  
XX Claim 30; Fig 16; 301pp; English.  
XX  
XX The invention relates to a purified polypeptide with alpha-amylase  
XX activity and the polynucleotide encoding it. The polypeptide is useful  
XX for hydrolysing starch linkages, for catalysing the breakdown of a  
XX starch, for modifying small molecules, for liquefying starch, for washing  
XX an object, for textile desizing, for treating lignocellulosic fibers, for  
XX improving fibre properties, for enzymatic de-linking of recycled paper  
XX pulp, for producing a high-maltose or high-glucose syrup or a mixed  
XX syrup, and for increasing the flow of production fluids from a  
XX subterranean formation by removing a viscous, starch-containing, damaging  
XX fluid formed during production operations and found within the  
XX subterranean formation which surrounds a completed well bore. The  
XX polypeptide is also useful for preparing a dough or a baked product  
XX prepared from the dough and in corn-wet milling processes, detergents,  
XX baking processes, beverages, oil fields (fuel ethanol), brewing processes  
XX and starch modification in the paper and pulp industry, for removing  
XX starch containing stains from a material and for reducing staling of  
XX bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
XX polypeptides of the invention  
XX  
SQ Sequence 436 AA;

Query Match 96.7%; Score 2378; DB 6; Length 436;  
Best Local Similarity 96.1%; Pred. No. 3.1e-196;  
Matches 419; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAKYSELEGGVIMQAFYNDVPSSGGIWDITIRQKIPEDYDAGISAIWIIPASKMGSGAYS 60  
DB 1 MAKYSELEGGVIMQAFYNDVPSSGGIWDITIRQKIPEDYDAGISAIWIIPASKMGSGAYS 60  
QY 61 MGYPDPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
DB 61 MGYPDPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
QY 121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180  
DB 121 NPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSDQYWLWAS 180  
QY 181 QESYAALYRISIGIDAWRFDYKGYAPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
DB 181 QESYAALYRISIGIDAWRFDYKGYAPWVKDMLNMGWAVGEYDNTVDVNLWAYSSG 240  
QY 241 AKVPDFALYKMDAEFNNKIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
DB 241 AKVPDFALYKMDAEFNNKIPALVSALONGQTVVSRDPFKAVTFVANHDTDIWNKYPA 300  
QY 301 YAFILTYEGQPTIFRYDYEELNKKLNIWHENLAGSGTDIVYDNDDELIFVRNGYG 360  
DB 301 YAFILTYEGQPTIFRYDYEELNKKLNIWHENLAGSGTDIVYDNDDELIFVRNGYG 360  
QY 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
DB 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDP 420

QY 421 ANGOYGYVWSYCGVG 436  
 DB 421 ANGOYGYVWSYCGVG 436

RESULT 6  
 ABU03057  
 ID ABU03057 standard; protein; 436 AA.  
 AC ABU03057;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Alpha amylase polypeptide #18.  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX  
 OS Unidentified.  
 XX  
 PN WC200268589-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 21-FEB-2002; 2002WO-US005068.  
 XX  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX  
 PA (DIVE-) DIVERSA CORP.  
 XX  
 PI Callen W, Richardson T, Frey G;  
 DR WPI; 2003-018656/01.  
 DR N-PSDB; ABX08428.  
 XX  
 PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX  
 PS Claim 30; Fig 16; 301pp; English.  
 XX  
 CC The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX  
 SQ Sequence 436 AA;

Query Match 96.7%; Score 2377; DB 6; Length 436;  
 Best Local Similarity 96.8%; Pred. No. 3.8e-196;  
 Matches 422; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWMDDTIROKIPWYDAGISAIWIPPASKMGGAYS 60  
 DB 1 MAKYLELEEGGVIMQAFYWDVPSGGIWMDDTIROKIPWYDAGISAIWIPPASKMGGAYS 60  
 QY 61 MGYDPYDFDLEGEYDOKGTETREFGSKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120  
 DB 61 MGYDPYDFDLEGEYDOKGTETREFGSKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120  
 QY 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNELHAGDSGTGGYPDI CHDKSWDOYWLWAS 180  
 DB 121 NPFVNDYTWTFDSKVASGKYTANYLDFHPNELHAGDSGTGGYPDI CHDKSWDOYWLWAS 180  
 QY 181 QESYAAVILRSIGIDAWRFDYKGYAPVWVKDNLNMGWAGVGEYWDNTVDALLNWAYSSG 240  
 DB 181 QESYAAVILRSIGIDAWRFDYKGYAPVWVKDNLNMGWAGVGEYWDNTVDALLNWAYSSG 240  
 QY 241 AKVFDFALYKMDAFDNKNI PALVSALONGQTQVSRDPKAVTFVANHTDIIWNKYPA 300  
 DB 241 AKVFDFPLYYKMDAFDNKNI PALVSALONGQTQVSRDPKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGQPTIPYRDYEELNKKLKNLIWIHNLNLAGGSTDIVYDNDDELIFVRNGY 360  
 DB 301 YAFILTYEGQPTIPYRDYEELNKKLKNLIWIHNLNLAGGSTDIVYDNDDELIFVRNGY 360  
 QY 361 DKPLITVINGSGSKAGRWVVPKEAGACIHEYTGNLGMWVKYVYSSGNVYLEAPAYDP 420  
 DB 361 SKPLITVINGSGSKAGRWVVPKEAGACIHEYTGNLGMWVKYVYSSGNVYLEAPAYDP 420  
 QY 421 ANGOYGYVWSYCGVG 436  
 DB 421 ANGOYGYVWSYCGVG 436

RESULT 7  
 ABU03053  
 ID ABU03053 standard; protein; 436 AA.  
 AC ABU03053;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Alpha amylase polypeptide #14.  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX  
 OS Unidentified.  
 XX  
 PN WC200268589-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 21-FEB-2002; 2002WO-US005068.  
 XX  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX  
 PA (DIVE-) DIVERSA CORP.  
 XX  
 PI Callen W, Richardson T, Frey G;  
 DR WPI; 2003-018656/01.  
 DR N-PSDB; ABX08424.  
 XX  
 PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.

high glucose syrup; corn-wet milling process; detergent; baking process; beverage; oil field; fuel ethanol; brewing process; stalling; starch modification.

Unidentified.

W0200268589-A2.

06-SEP-2002.

21-FEB-2002; 2002WO-US005068.

21-FEB-2001; 2001US-0270495P.

21-FEB-2001; 2001US-0270496P.

14-MAY-2001; 2001US-0291122P.

(DIVE-) DIVERSA CORP.

Callen W, Richardson T, Frey G;

WPI; 2003-018656/01.

N-FSD3; ABX08412.

Novel purified polypeptide with alpha-amylase activity, useful e.g. for liquefying starch, for textile desizing, for treating lignocellulosic fibers, and for producing high-maltose or high-glucose syrup.

Claim 30; Fig 16; 301pp; English.

The invention relates to a purified polypeptide with alpha-amylase activity and the polynucleotide encoding it. The polypeptide is useful for hydrolysing starch linkages, for catalysing the breakdown of a starch, for modifying small molecules, for liquefying starch, for washing an object, for textile desizing, for treating lignocellulosic fibers, for improving fibre properties, for enzymatic de-inking of recycled paper pulp, for producing a high-maltose or high-glucose syrup or a mixed syrup, and for increasing the flow of production fluids from a subterranean formation by removing a viscous, starch-containing, damaging fluid formed during production operations and found within the subterranean formation which surrounds a completed well bore. The polypeptide is also useful for preparing a dough or a baked product prepared from the dough and in corn-wet milling processes, detergents, baking processes, beverages, oil fields (fuel ethanol), brewing processes and starch modification in the paper and pulp industry, for removing starch containing stains from a material and for reducing staling of bakery products. Sequences ABU03040-ABU03144 represent alpha amylase polypeptides of the invention

Sequence 436 AA;

Query Match

Best Local Similarity

Matches 418; Conservative

Score 2371; DB 6; Length 436;

Pred. No. 1.2e-195;

Mismatches 10; Indels 0; Gaps 0;

1 MAKYSELEKGVIMQAFYWDVPSGGTWTDTIRQKIPEDWDAGISAIWIPPSKMGCGAYS 60

1 MAKYLEEGGLIMQAFYWDVPMGGIWDVTVAQKIPDMASAGISAIWIPPSKMGCGAYS 60

61 MGYDPPDFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVADIIVNHRAGGDLW 120

61 MGYDPPDFDLGEBYDQKGTETRFSGSKOELVNMINTAHAYGMKVADIIVNHRAGGDLW 120

121 NPFVNDYTWDTDFSKVAGSKYTANYLDFHPNELHAGDSGTFGGYPDI:CHDKSWDQYWLWAS 180

121 NPFVNDYTWDTDFSKVAGSKYTANYLDFHPNELHAGDSGTFGGYPDI:CHDKSWDQYWLWAS 180

181 QESYAALYLRISIGIDAWRFDYVKGAPVWVYKDWLNWGWGAVGEYWDVTNVDVNLWYSSG 240

181 QESYAALYLRISIGIDAWRFDYVKGAPVWVYKDWLNWGWGAVGEYWDVTNVDVNLWYSSG 240

241 AKVDFPDLALYKMDAEAFDNKNI:PALVSALQNGQTVVSRDPPFKAVTFVANHTDIIWNKYPA 300

241 AKVDFPDLALYKMDAEAFDNKNI:PALVSALQNGQTVVSRDPPFKAVTFVANHTDIIWNKYLA 300

301 YAFILTYEGQTFIPYRDYEEWLNKDKLNLIWIHENLAGGSTDIVYDNDDELIFVRNGYG 360

301 YAFILTYEGQTFIPYRDYEEWLNKDKLNLIWIHENLAGGSTDIVYDNDDELIFVRNGYG 360

361 DKPLIYIYILNGSKAGRWVYVPRKAGAC:THEYTNLNGWVDKYVYSSGWVYLEAPAYDP 420

361 DKPLIYIYILNGSKAGRWVYVPRKAGAC:THEYTNLNGWVDKYVYSSGWVYLEAPAYDP 420

421 ANGOYGYSVMSYCGVG 436

421 ANGOYGYSVMSYCGVG 436

RESULT 8

ABU03041

ID ABU03041 standard; protein; 436 AA.

XX AC ABU03041;

XX DT 21-JAN-2003 (first entry)

XX DE Alpha amylase polypeptide #2.

XX KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;

XX KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;

XX KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;

XX KW

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 301 YAFILTYEGOPTIFRYDYBEWLNKDKLNLWIHNLNLAGGSTDIVYDNDDELIFVRNGYG 360  
 Db |||||  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEYTGNLGGWDXVYVSSGWVLEAPADPP 420  
 Db |||||  
 QY 421 ANQGYGYSVMSYCGVG 436  
 Db |||||

RESULT 9  
 ABU03058  
 ID ABU03058 standard; protein; 436 AA.  
 AC  
 DT 21-JAN-2003 (first entry)  
 DE Alpha amylase polypeptide #19.  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 OS Unidentified.  
 PN WO200268589-A2.  
 PD 06-SEP-2002.  
 PF 21-FEB-2002; 2002WO-US005068.  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 PA Callen W, Richardson T, Frey G;  
 PI WPI: 2003-018556/01.  
 DR N-PSDB; ABX08429.

Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 liquefying starch, for textile desizing, for treating lignocellulosic  
 fibers, and for producing high-maltose or high-glucose syrup.

Claim 30; Fig 16; 30lpp; English.

The invention relates to a purified polypeptide with alpha-amylase  
 activity and the polynucleotide encoding it. The polypeptide is useful  
 for hydrolysing starch linkages, for catalysing the breakdown of a  
 starch for modifying small molecules, for liquefying starch, for washing  
 an object for textile desizing, for treating lignocellulosic fibers, for  
 improving fibre properties, for enzymatic de-linking of recycled paper  
 pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 syrup, and for increasing the flow of production fluids from a  
 subterranean formation by removing a viscous, starch-containing, damaging  
 fluid formed during production operations and found within the  
 subterranean formation which surrounds a completed well bore. The  
 polypeptide is also useful for preparing a dough or a baked product  
 prepared from the dough and in corn-wet milling processes, detergents,  
 baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 and starch modification in the paper and pulp industry, for removing  
 starch containing stains from a material and for reducing staling of  
 bakery products. Sequences ABU03040-ABU03144 represent alpha amylase

CC polypeptides of the invention  
 XX  
 SQ Sequence 436 AA;  
 Query Match 96.0%; Score 2360; DB 6; Length 436;  
 Best Local Similarity 95.6%; Pred. No. 1.1e-194;  
 Matches 417; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 1 MAKYSELEKGVINQAFVWDPVSGGIWMDTIROKIPWYDAGISAIWIPASKMGSGAYS 60  
 Db |||||  
 QY 61 MGYPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYCMKVIADIVINHRAGGLEW 120  
 Db |||||  
 QY 121 NPFVNDYTWTFDSKVASGKYTANYLDLPHNELHAGDSGTGGYPDICHDKSWOYMLWAS 180  
 Db |||||  
 QY 181 QESYAAVLRISIGIDANWFDYVKGYPWVVKDNLNMGWAVGEYWDNTVDVNLWAYSSG 240  
 Db |||||  
 QY 241 AKVDFPALLYKMDAEPDNKNI PALVSALONGQTWVSQRPKAVTFVANHTDIIWNKYPA 300  
 Db |||||  
 QY 301 YAFILTYEGOPTIFRYDYBEWLNKDKLNLWIHNLNLAGGSTDIVYDNDDELIFVRNGYG 360  
 Db |||||  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKFAACIHEYTGNLGGWDXVYVSSGWVLEAPADPP 420  
 Db |||||  
 QY 421 ANQGYGYSVMSYCGVG 436  
 Db |||||

RESULT 10  
 ABU03046  
 ID ABU03046 standard; protein; 436 AA.  
 AC ABU03046;  
 DT 21-JAN-2003 (first entry)  
 DE Alpha amylase polypeptide #7.  
 KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 OS Unidentified.  
 PN WO200268589-A2.  
 PD 06-SEP-2002.  
 PF 21-FEB-2002; 2002WO-US005068.  
 PR 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 PA Callen W, Richardson T, Frey G;



XX WPI; 2003-018656/01.  
 DR N-PSDB; ABX08417.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 30; Fig 16; 301pp; English.  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX Sequence 436 AA;  
 SQ  
 Query Match 95.9%; Score 2358; DB 6; Length 436;  
 Best Local Similarity 95.4%; Pred. No. 1.6e-194;  
 Matches 416; Conservative 9; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 DB 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 QY 61 MGVDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 DB 61 MGVDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTDWTFPSKVASGKTYANLDFHPNHLHAGSDGTFGGYPDICHDKSWDQYWLWAS 180  
 DB 121 NPFVNDYTDWTFPSKVASGKTYANLDFHPNHLHAGSDGTFGGYPDICHDKSWDQYWLWAS 180  
 QY 181 QESYAAYLRISIGIDAWRFYVKGYPVWVVDLNMWGGWAVGEYWDVTNVDALNWAYSSG 240  
 DB 181 QESYAAYLRISIGIDAWRFYVKGYPVWVVDLNMWGGWAVGEYWDVTNVDALNWAYSSD 240  
 QY 241 AKVFDFALYKMDFAFNKNIIPALVSALONGQVTWSDRDPKATTFVANHDTDIWNKYPA 300  
 DB 241 AKVFDFALYKMDFAFNKNIIPALVSALONGQVTWSDRDPKATTFVANHDTDIWNKYPA 300  
 QY 301 YAFLLTYEGOPTIFRYDYEWLNKDKLNLITHEINLAGGSTDIVYDNDDELIFVNGYG 360  
 DB 301 YAFLLTYEGOPTIFRYDYEWLNKDKLNLITHEINLAGGSTDIVYDNDDELIFVNGYG 360  
 QY 361 DKPLGITYINLSSKAGRWVVPKFAACIHEYTGNLGGWVDKYVYSSGAVYLEAPAYDP 420  
 DB 361 SKPLGITYINLSSKAGRWVVPKFAACIHEYTGNLGGWVDKYVYSSGAVYLEAPAYDP 420  
 QY 421 ANGOYGYSVNSYCGVG 436  
 DB 421 ANGOYGYSVNSYCGVG 436  
 RESULT 11  
 ABU03069  
 ID ABU03069 standard; protein; 436 AA.  
 XX  
 AC ABU03069;

XX 21-JAN-2003 (first entry)  
 XX Alpha amylase polypeptide #30.  
 DE  
 XX Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 XX Unidentified.  
 OS  
 XX WO200269589-A2.  
 FN  
 XX 06-SEP-2002.  
 PD  
 XX 21-FEB-2002; 2002WO-US005068.  
 PF  
 XX 21-FEB-2001; 2001US-0270495P.  
 PR  
 XX 21-FEB-2001; 2001US-0270496P.  
 PR  
 XX 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 PA  
 XX Callen W, Richardson T, Frey G;  
 PI  
 XX WPI; 2003-018656/01.  
 DR  
 XX N-PSDB; ABX08440.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 30; Fig 16; 301pp; English.  
 PS The invention relates to a purified polypeptide with alpha-amylase  
 XX activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX Sequence 436 AA;  
 SQ  
 Query Match 95.9%; Score 2358; DB 6; Length 436;  
 Best Local Similarity 95.4%; Pred. No. 1.6e-194;  
 Matches 416; Conservative 11; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 DB 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIIPASKMGWGAYS 60  
 QY 61 MGVDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 DB 61 MGVDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTDWTFPSKVASGKTYANLDFHPNHLHAGSDGTFGGYPDICHDKSWDQYWLWAS 180  
 DB 121 NPFVNDYTDWTFPSKVASGKTYANLDFHPNHLHAGSDGTFGGYPDICHDKSWDQYWLWAS 180  
 QY 181 QESYAAYLRISIGIDAWRFYVKGYPVWVVDLNMWGGWAVGEYWDVTNVDALNWAYSSG 240  
 DB 181 QESYAAYLRISIGIDAWRFYVKGYPVWVVDLNMWGGWAVGEYWDVTNVDALNWAYSSD 240  
 QY 241 AKVFDFALYKMDFAFNKNIIPALVSALONGQVTWSDRDPKATTFVANHDTDIWNKYPA 300  
 DB 241 AKVFDFALYKMDFAFNKNIIPALVSALONGQVTWSDRDPKATTFVANHDTDIWNKYPA 300  
 QY 301 YAFLLTYEGOPTIFRYDYEWLNKDKLNLITHEINLAGGSTDIVYDNDDELIFVNGYG 360  
 DB 301 YAFLLTYEGOPTIFRYDYEWLNKDKLNLITHEINLAGGSTDIVYDNDDELIFVNGYG 360  
 QY 361 DKPLGITYINLSSKAGRWVVPKFAACIHEYTGNLGGWVDKYVYSSGAVYLEAPAYDP 420  
 DB 361 SKPLGITYINLSSKAGRWVVPKFAACIHEYTGNLGGWVDKYVYSSGAVYLEAPAYDP 420  
 QY 421 ANGOYGYSVNSYCGVG 436  
 DB 421 ANGOYGYSVNSYCGVG 436

QY 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWWGGWAVGEYDNTNVDVNLWAYSSG 240  
DB 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWWGGWAVGEYDNTNVDVNLWAYSSD 240  
QY 241 AKVDFALYKMDAEAFDNKNIIPALVSAALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
DB 241 AKVDFDPLYYKMDAEAFDNKNIIPALVSAALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
QY 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
DB 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
QY 361 DKPGLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
DB 361 DREPLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPADP 420  
QY 421 ANGGYGSVWSYCGVG 436  
DB 421 ANGGYGSVWSYCGVG 436  
RESULT 12  
ABU03054  
ID ABU03054 standard; protein; 436 AA.  
AC ABU03054;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Alpha amylase polypeptide #15.  
XX  
KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
OS Unidentified.  
XX  
PN WO200268589-A2.  
XX  
PD 06-SEP-2002.  
XX  
XX 21-FEB-2002; 2002WO-US005068.  
XX  
XX 21-FEB-2001; 2001US-0270495P.  
PR 21-FEB-2001; 2001US-0270495P.  
PR 14-MAY-2001; 2001US-0291122P.  
XX  
PA (DIVE-) DIVERSA CORP.  
XX  
PI Callen W, Richardson T, Frey G;  
XX  
DR WPI; 2003-018656/01.  
DR N-PSDE; ABX08425.  
XX  
XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
PT liquefying starch, for textile desizing, for treating lignocellulosic  
PT fibers, and for producing high-maltose or high-glucose syrup.  
XX  
PS Claim 30; Fig 16; 301pp; English.  
XX  
CC The invention relates to a purified polypeptide with alpha-amylase  
CC activity and the polynucleotide encoding it. The polypeptide is useful  
CC for hydrolysing starch linkages, for catalysing the breakdown of a  
CC starch, for modifying small molecules, for liquefying starch, for washing  
CC an object, for textile desizing, for treating lignocellulosic fibers, for  
CC improving fibre properties, for enzymatic de-linking of recycled paper  
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
CC syrup, and for increasing the flow of production fluids from a  
CC subterranean formation by removing a viscous, starch-containing, damaging

CC fluid formed during production operations and found within the  
CC subterranean formation which surrounds a completed well bore. The  
CC polypeptide is also useful for preparing a dough or a baked product  
CC prepared from the dough and in corn-wet milling processes, detergents,  
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
CC and starch modification in the paper and pulp industry, for removing  
CC starch containing stains from a material and for reducing staling of  
CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
CC polypeptides of the invention  
XX  
SQ Sequence 436 AA;  
Query Match 95.8%; Score 2356; DB 6; Length 436;  
Best Local Similarity 95.0%; Pred. No. 2.4e-194;  
Matches 414; Conservative 9; Mismatches 13; Indels 0; Gaps 0;  
QY 1 MAKYSELEKGVINQAFYWDVPSGGIWMWDTIRKIPEDYDAGISAIWIPASKMGGAYS 60  
DB 1 MAKYSELEKGVINQAFYWDVPSGGIWMWDTIRKIPEDYDAGISAIWIPASKMGGAYS 60  
QY 61 MGYPDYDFDLGEYDQKGTETFRFGSKQELVNMINTAHAYGKVIADIVINHRAGGLEW 120  
DB 61 MGYPDYDFDLGEYDQKGTETFRFGSKQELVNMINTAHAYGKVIADIVINHRAGGLEW 120  
QY 121 NPFVNDYTWTDKSVASGKYTANYLDPHNLHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
DB 121 NPFVNDYTWTDKSVASGKYTANYLDPHNLHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
QY 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWWGGWAVGEYDNTNVDVNLWAYSSG 240  
DB 181 QESYAAALRSIGIDAWRFDYVKGAPWVVKDLNWWGGWAVGEYDNTNVDVNLWAYSSG 240  
QY 241 AKVDFALYKMDAEAFDNKNIIPALVSAALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
DB 241 AKVDFDPLYYKMDAEAFDNKNIIPALVSAALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300  
QY 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
DB 301 YAFILTYEGOPTIFYRDYBEWLNKDKLNIWIHENLAGGSTDIVYDNDLIFVRNGYG 360  
QY 361 DKPGLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
DB 361 SKPGLITYINLGSSKAGRWVYVVKPFAGACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420  
QY 421 ANGGYGSVWSYCGVG 436  
DB 421 ANGGYGSVWSYCGVG 436  
RESULT 13  
ABU03048  
ID ABU03048 standard; protein; 436 AA.  
XX  
AC ABU03048;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Alpha amylase polypeptide #9.  
XX  
KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
KW high glucose syrup; corn-wet milling process; detergent; baking process;  
KW beverage; oil field; fuel ethanol; brewing process; staling;  
KW starch modification.  
XX  
OS Unidentified.  
XX  
XX WO200268589-A2.  
XX  
PD 06-SEP-2002.  
XX  
XX 21-FEB-2002; 2002WO-US005068.

XX 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 XX Callen W, Richardson T, Frey G;  
 DR WPI; 2003-018656/01.  
 DR N-PSDB; ABX08419.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 30; Fig 16; 301pp; English.  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX Sequence 436 AA;  
 XX

Query Match 95.8%; Score 2356; DB 6; Length 436;  
 Best Local Similarity 95.2%; Pred. No. 2.4e-194;  
 Matches 415; Conservative 10; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 DB 1 MAKYLEEGGLIQAFAFDVPMGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 QY 61 MGYPDYFDLGEYDQGTVEFRSGKQELVNMINTAHAYCMKVIADIVINHRAGDLEW 120  
 DB 61 MGYPDYFDLGEYDQGTVEFRSGKQELVNMINTAHAYCMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYDFTDFSVASGKYTANYLDHPHNLHAGSGTGGYDICHKSDQYWLWAS 180  
 DB 121 NPFVGDYDFTDFSVASGKYTANYLDHPHNLHAGSGTGGYDICHKSDQYWLWAS 180  
 QY 181 QESYAALRSIGIDAWRFDYKGYAPVWVVDLNLWNGWAGVGYWDTNVDVNLWAYSSG 240  
 DB 181 QESYAALRSIGIDAWRFDYKGYAPVWVVDLNLWNGWAGVGYWDTNVDVNLWAYSSG 240  
 QY 241 AKVDFPALYKWDFAFNKNTIPALVSALONGQTVSRDPEKATVFVANHDTDIWNKYPA 300  
 DB 241 AKVDFPALYKWDFAFNKNTIPALVSALONGQTVSRDPEKATVFVANHDTDIWNKYPA 300  
 QY 301 YAFILTYEGQPTIFRYDYEWLNKDKLNLIWIHNLNLAGGSTDIVYDNDLIFVRNGYG 360  
 DB 301 YAFILTYEGQPTIFRYDYEWLNKDKLNLIWIHNLNLAGGSTDIVYDNDLIFVRNGYG 360  
 QY 361 DKPGLITVINLGSAGRWVYVPKFAACIHEVTGNLGGWVDKYVYSSGWVILEAPAYDP 420  
 DB 361 SKPGLITVINLGSAGRWVYVPKFAACIHEVTGNLGGWVDKYVYSSGWVILEAPAYDP 420  
 QY 421 ANGQGYSVWSYCGVG 436  
 DB 1 MAKYLEEGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 DB 1 MAKYLEEGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60

Db 421 ANGQGYSVWSYCGVG 436

RESULT 14  
 ABU03066  
 ID ABU03066 standard; protein; 436 AA.  
 XX AC ABU03066;  
 XX 21-JAN-2003 (first entry)  
 XX Alpha amylase polypeptide #27.  
 XX Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; staling;  
 KW starch modification.  
 KW Unidentified.  
 OS WO200268589-A2.  
 XX 06-SEP-2002.  
 XX 21-FEB-2002; 2002WO-US005068.  
 XX 21-FEB-2001; 2001US-0270495P.  
 PR 21-FEB-2001; 2001US-0270496P.  
 PR 14-MAY-2001; 2001US-0291122P.  
 XX (DIVE-) DIVERSA CORP.  
 XX Callen W, Richardson T, Frey G;  
 DR WPI; 2003-018656/01.  
 DR N-PSDB; ABX08437.  
 XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for  
 PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.  
 XX Claim 30; Fig 16; 301pp; English.  
 XX The invention relates to a purified polypeptide with alpha-amylase  
 CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX Sequence 436 AA;  
 XX

Query Match 95.6%; Score 2352; DB 6; Length 436;  
 Best Local Similarity 95.2%; Pred. No. 5.4e-194;  
 Matches 415; Conservative 10; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 DB 1 MAKYLEEGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60

Sequence 436 AA;  
 Query Match 95.6%; Score 2352; DB 6; Length 436;  
 Best Local Similarity 95.2%; Pred. No. 5.4e-194;  
 Matches 415; Conservative 10; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60  
 DB 1 MAKYLEEGGVIMQAFYWDVPSGGIWDITROKIPEDWDAGISAIWIPASKMGGAYS 60

QY 61 MGYPDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 DB 61 MGYPDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWDFSKVASGKTYANTYLDHFNELHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
 DB 121 NPFVNDYTWDFSKVASGKTYANTYLDHFNELHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
 QY 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYWDNTVDVNLWAYSSG 240  
 DB 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYWDNTVDVNLWAYSSG 240  
 QY 241 AKVDFDALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 DB 241 AKVDFDALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSSTDIYYDNDDELIFVRNGYG 360  
 DB 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSSTDIYYDNDDELIFVRNGYG 360  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKAGACIHEVTGNLGGWVDKYVYSSGWWYLEAPAYDP 420  
 DB 361 SKPGLITYINLGSSKAGRWVYVPKAGACIHEVTGNLGGWVDKYVYSSGWWYLEAPAYDP 420  
 QY 421 ANGQYGSVMSYCGVG 436  
 DB 421 ANGQYGSVMSYCGVG 436

## RESULT 15

ABU03043  
 ID ABU03043 standard; protein; 436 AA.

AC ABU03043;  
 XX

DT 21-JAN-2003 (first entry)

DE Alpha amylase polypeptide #4.

KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;  
 KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;  
 KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;  
 KW high glucose syrup; corn-wet milling process; detergent; baking process;  
 KW beverage; oil field; fuel ethanol; brewing process; scaling;  
 KW starch modification.

OS Unidentified.

PN WO200268589-A2.

PD 06-SEP-2002.

PF 21-FEB-2002; 2002WO-US005068.

PR 21-FEB-2001; 2001US-0270495P.

PR 21-FEB-2001; 2001US-0270496P.

PR 14-MAY-2001; 2001US-0291122P.

PA (DIVE-) DIVERSA CORP.

XX Callen W, Richardson T, Frey G;

XX WPI; 2003-018656/01.

DR N-PSDB; ABX08414.

XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for

PT liquefying starch, for textile desizing, for treating lignocellulosic  
 PT fibers, and for producing high-maltose or high-glucose syrup.

PS Claim 30; Fig 16; 301pp; English.

XX The invention relates to a purified polypeptide with alpha-amylase

CC activity and the polynucleotide encoding it. The polypeptide is useful  
 CC for hydrolysing starch linkages, for catalysing the breakdown of a  
 CC starch, for modifying small molecules, for liquefying starch, for washing  
 CC an object, for textile desizing, for treating lignocellulosic fibers, for  
 CC improving fibre properties, for enzymatic de-linking of recycled paper  
 CC pulp, for producing a high-maltose or high-glucose syrup or a mixed  
 CC syrup, and for increasing the flow of production fluids from a  
 CC subterranean formation by removing a viscous, starch-containing, damaging  
 CC fluid formed during production operations and found within the  
 CC subterranean formation which surrounds a completed well bore. The  
 CC polypeptide is also useful for preparing a dough or a baked product  
 CC prepared from the dough and in corn-wet milling processes, detergents,  
 CC baking processes, beverages, oil fields (fuel ethanol), brewing processes  
 CC and starch modification in the paper and pulp industry, for removing  
 CC starch containing stains from a material and for reducing staling of  
 CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase  
 CC polypeptides of the invention  
 XX  
 SQ Sequence 436 AA;

Query Match 95.6%; Score 2350; DB 6; Length 436;  
 Best Local Similarity 95.2%; Pred. No. 8e-194;  
 Matches 415; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 MAKYSELEKGGVIMQAFYWDVPSGGIWWDTIROKIPEDYDAGISAIWIPPSKGMGGAYS 60  
 DB 1 MAKYSELEKGGVIMQAFYWDVPSGGIWWDTIROKIPEDYDAGISAIWIPPSKGMGGAYS 60  
 QY 61 MGYPDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 DB 61 MGYPDYDFDLGEYDQKGTETFRGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWDFSKVASGKTYANTYLDHFNELHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
 DB 121 NPFVNDYTWDFSKVASGKTYANTYLDHFNELHAGDSGTGGYPDICHDKSWDQYWLWAS 180  
 QY 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYWDNTVDVNLWAYSSG 240  
 DB 181 QESYAAYLRSIGIDAWRFDYVKGYPWVKDMLNMGWAVGEYWDNTVDVNLWAYSSG 240  
 QY 241 AKVDFDALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 DB 241 AKVDFDALYKMDAEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSSTDIYYDNDDELIFVRNGYG 360  
 DB 301 YAFILTYEGOPTIFYRDYEELNKKDLKNIWIHENLAGSSTDIYYDNDDELIFVRNGYG 360  
 QY 361 DKPGLITYINLGSSKAGRWVYVPKAGACIHEVTGNLGGWVDKYVYSSGWWYLEAPAYDP 420  
 DB 361 SKPGLITYINLGSSKAGRWVYVPKAGACIHEVTGNLGGWVDKYVYSSGWWYLEAPAYDP 420  
 QY 421 ANGQYGSVMSYCGVG 436  
 DB 421 ANGQYGSVMSYCGVG 436

Search completed: June 29, 2004, 09:34:14  
 Job time : 61 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:32:02 ; Search time 21 Seconds  
(without alignments)  
1997.119 Million cell updates/sec

Title: US-10-081-739A-2

Perfect score: 2459

Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQYGVSVKSGVG 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	492	20.0	492	2 AH2079	alpha-amylase (imp
2	487.5	19.8	494	2 AD0751	cytoplasmic alpha-
3	486.5	19.8	494	1 B45738	alpha-amylase (EC
4	475.5	19.3	549	1 A24549	alpha-amylase (EC
5	472.5	19.2	514	1 A24541	alpha-amylase (EC
6	466.5	19.0	549	1 A24541	alpha-amylase (EC
7	461.5	18.8	549	1 A24436	alpha-amylase (EC
8	459	18.7	548	1 A24541	alpha-amylase (EC
9	458.5	18.6	512	1 A24541	alpha-amylase (EC
10	454.5	18.5	518	1 A27705	alpha-amylase (EC
11	452.5	18.4	495	2 B90962	cytoplasmic alpha-
12	451.5	18.4	495	2 AD3038	alpha-amylase (EC
13	451.5	18.4	506	2 G98247	alpha-amylase (EC
14	447.5	18.2	495	1 A25738	cytoplasmic alpha-
15	444.5	18.1	495	2 B5810	alpha-amylase (EC
16	438	17.8	491	2 C86781	alpha-amylase (EC
17	420	17.1	493	2 S10514	alpha-amylase (EC
18	400	16.3	421	2 T09942	alpha-amylase (EC
19	398	16.2	423	2 J01527	alpha-amylase (EC
20	396.5	16.1	443	2 S19990	alpha-amylase (EC
21	390.5	15.9	445	2 S19990	alpha-amylase (EC
22	373	15.2	484	2 G95160	alpha-amylase (imp
23	372	15.1	484	2 F98026	alpha-amylase (EC
24	366.5	14.9	427	1 ALBHB	alpha-amylase (EC
25	365.5	14.9	438	1 ALBHB	alpha-amylase (EC
26	363.5	14.8	504	2 A55861	alpha-amylase (EC
27	357.5	14.5	438	2 S14957	alpha-amylase (EC
28	353	14.4	437	2 S14956	alpha-amylase (EC
29	351.5	14.3	429	1 J50406	alpha-amylase (EC

30 348.5 14.2 428 2 T05521 alpha-amylase (EC  
31 347.5 14.1 439 2 T02956 alpha-amylase (EC  
32 347.5 14.1 440 2 S14958 alpha-amylase (EC  
33 347 14.1 428 2 S10013 alpha-amylase (EC  
34 347 14.1 434 2 S12775 alpha-amylase (EC  
35 341.5 13.9 435 2 S12625 alpha-amylase (EC  
36 341.5 13.9 435 2 JC7137 alpha-amylase (EC  
37 340.5 13.8 437 2 S07040 alpha-amylase (EC  
38 339 13.8 826 2 E96720 probable alpha-amy  
39 336 13.7 437 2 JT0946 alpha-amylase (EC  
40 336 13.7 437 2 JC7138 alpha-amylase (EC  
41 335.5 13.6 413 1 ALWT3 alpha-amylase (EC  
42 294 12.0 412 2 C96789 protein T23E18.6  
43 288.5 11.7 547 2 A32803 Glucan 1,4-alpha-m  
44 278 11.3 551 2 S05667 glucan 1,4-alpha-m  
45 253.5 10.3 710 2 S63598 cyclomaltodextrin

#### ALIGNMENTS

##### RESULT 1

AH2079  
alpha-amylase [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AH2079  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anai  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2079  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-492 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA73889.1; PID:gl7131281; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr2190  
C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology

Query Match 20.0%; Score 492; DB 2; Length 492;  
Best Local Similarity 27.8%; Pred. No. 4.6e-28;  
Matches 147; Conservative 81; Mismatches 153; Indels 148; Gaps 28;  
QY 8 EKGGVIMQAFYWDVPSGGIWDTIROKIPWYDAGISAIWIPPSKMGCGAYSMGYDPTD 67  
Db 3 QMGTMMQVPHWYIPNDGNLWSKVEASAPELADAGFTAMMLPPAYKGFAGSFDVGYGYD 62  
QY 68 FFDLGEYDQGTVEFRFGSKQELVNNINTAHAYGMKVIADIVINHAGDGL---EWNPF 123  
Db 63 LFDLGEFDQKGSVRTKYGTFRQYLDVAVSLQTHGLQVADAVLNHNKMGSDAVETPKATPF 122  
QY 124 VND-----YTWDPSKVASGYTANY-----LDHFNELHAGDSGT- 159  
Db 123 PQDRLNPKGGLQDIKTYTHYFPFG-RQGY-SNFEWHWHHFDVGY--NEVNSGDRSTV 178  
QY 160 -----FGYYP-----DICHKSKNDQ---YW-LWASQESVAAVLRISIGD 194  
Db 179 YLLEGNFDDYVALEKGNFAYLMGCDLDFQNEWRVREVTYWGKWC-----LDTTKVD 230  
QY 195 AMFEDYKVGAPVWVKDNLNWMGGWA-----VGEYWDNTVDVAVLNW---AYSSGAKVFD 245  
Db 231 GFRIDAIKHISTWFFPEWIDALERHAGKOLFVGEYWNIDINTLL-WYWDVAVRGMKSVFD 289  
QY 246 FALYYKMDA-----FDKNKIPALVSALQNGVTVSVRDPFKAVTFVANHDTII----- 294  
Db 290 VPLHYNFHQASGSGNIDMRRI-----LDG-TMMQQRPTAVTFVENHDSOPQALLES 341  
QY 295 ----WNKYPAYAFI-LTYEGQPTIFYRD-----YEWNLNKDK-----LKNLI 331

Db 342 VVEPWFPLAYAILLRQEGYPCVFHADYGAIEDW-GKGNRYNIFPMSHRWIIDKLL 400  
Qy 332 WIHENLAGGSTDIVYDNDLI-FVRNGYDGPGLITVINGSSKAGRWVYV---PKFA 386  
Db 401 YARKHAYG-PQYVLDHNTTGTWLTGADHPQGMVIMDSGSGIKMVGKPTKFI 459  
Qy 387 GACIHEYTONLGGWDKYYVSGGWVLEAPAYDPANGQY-----GYSW 430  
Db 460 DLTEH-----IKEAVYTNEGW-----GBPRCLGSGSVW 489  
RESULT 2  
AD0751  
cytoplasmic alpha-amylase [imported] - Salmonella enterica subsp. enterica serovar Typhi  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 08-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AD0751  
R:Farrell, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AD0751  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-494 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD05711.1; PID:g16503204; GSPDB:GN00176  
C:Genetics:  
A:Gene: SY2171  
C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
Query Match 19.8%; Score 487.5; DB 2; Length 494;  
Best Local Similarity 28.7%; Pred. No. 9.9e-28;  
Matches 139; Conservative 71; Mismatches 156; Indels 119; Gaps 21;  
Qy 9 KGVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPSPASKMGAGYSGVDPYDF 68  
Db 2 KNPTLLQYFHWYDPGGKLSLAERADGLNDIGINWMLFPACKGASGYSVGYDYL 61  
Qy 69 FDLGEYDQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDGLWNPV 124  
Db 62 FDLGEYDQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDGLWNPV 121  
Qy 125 -NDYT-----WTDFSKVA-SGKYTANYLDHF-----PHELHAGDSGTF- 160  
Db 122 QDRTQIDDDNIIECEGWTRYTFPARAGQYSNFIWYHCFSGIDHIENPDE-----DGIFK 176  
Qy 161 -----GGYDICHDK--SWD-----QYWL-WASQESYAAYLSIG 192  
Db 177 IVNDYTGDNVDQVDDSLGNFDYLMGENIDFRNHAVTEELKYARWVMEQTHC----- 229  
Qy 193 IDAWRFYVKGYPWVVKMLNMWGGWA-----VGEYWDNTVDVNLWAYSSGAK--VF 244  
Db 230 -DGFRDLAVKHIPAWFYKEMIEHVAQVAPKPLFIVAEYWSHEVDKLTQYIDQADGKTM 288  
Qy 245 DFALYYKMDA-----FDKNKIPALVSALONGQTVVSRDPFKAVTFVANHDTII 294  
Db 289 DAPLQMKFHEASROGAEYDMRHIFT-----GTLVEADPFHATVILVANHDTIQLQALE 340  
Qy 295 -----WNKYPAYAFILTYE-QOPTIFYED-----EMLNKD-----KLNLIW 332  
Db 341 APVEPWFKPLAYALILLRENGVPSVFYDLYGASYEDNGENGETCRVDMPEVINGOLDRLIL 400  
Qy 333 IHENLAGGSTDIVYDNDLI-FVRNGYDGPGLITVINGSSKAGRWVYVVPFAGACIH 391  
Db 401 ARQRFAGH-IQTLFFDHPNCIAFSRSGTEENPGCVVLSNGDDGDKETLLGLDNYANKTW 459  
Qy 392 EYTG 396  
Db 460 DFLGN 464

RESULT 3  
B45738  
alpha-amylase (EC 3.2.1.1), cytosolic - Salmonella typhimurium  
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C:Species: Salmonella typhimurium  
C:Date: 07-Apr-1994 #sequence\_revision 18-Aug-1995 #text\_change 18-Jun-1999  
C:Accession: B45738  
R:Raha, M.; Kawagishi, I.; Mueller, V.; Kihara, M.; Macnab, R.M.  
J. Bacteriol. 174, 6644-6652, 1992  
A:Title: Escherichia coli produces a cytoplasmic alpha-amylase, amyA.  
A:Reference number: A45738; MUID:93015717; PMID:1400215  
A:Accession: B45738  
A:Molecule type: DNA  
A:Residues: 1-494 <RAH>  
A:Cross-references: GB:L01643; NID:g154043; PIDN:AAA27110.1; PID:g154045  
C:Genetics:  
A:Gene: amyA  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
C:Keywords: cytosol; glycosidase; hydrolase; polysaccharide degradation  
F:202-335/Domain: alpha-amylase core homology <AMY>  
F:239,265,332/Active site: His, Glu, Asp #status predicted  
Query Match 19.8%; Score 486.5; DB 1; Length 494;  
Best Local Similarity 28.2%; Pred. No. 1.2e-27;  
Matches 140; Conservative 75; Mismatches 159; Indels 123; Gaps 21;  
Qy 9 KGVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPSPASKMGAGYSGVDPYDF 68  
Db 2 KNPTLLQYFHWYDPGGKLSLAERADGLNDIGINWMLFPACKGASGYSVGYDYL 61  
Qy 69 FDLGEYDQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDGLWNPV 124  
Db 62 FDLGEYDQKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDGLWNPV 121  
Qy 125 -NDYT-----WTDFSKVA-SGKYTANYLDHF-----PHELHAGDSGTF- 160  
Db 122 QDRTQIDDDNIIECEGWTRYTFPARAGQYSNFIWYHCFSGIDHIENPDE-----DGIFK 176  
Qy 161 -----GGYDICHDK--SWD-----QYWL-WASQESYAAYLSIG 192  
Db 177 IVNDYTGDNVDQVDDSLGNFDYLMGENIDFRNHAVTEELKYARWVMEQTHC----- 229  
Qy 193 IDAWRFYVKGYPWVVKMLNMWGGWA-----VGEYWDNTVDVNLWAYSSGAK--VF 244  
Db 230 -DGFRDLAVKHIPAWFYKEMIEHVAQVAPKPLFIVAEYWSHEVDKLTQYIDQADGKTM 288  
Qy 245 DFALYYKMDA-----FDKNKIPALVSALONGQTVVSRDPFKAVTFVANHDTII 294  
Db 289 DAPLQMKFHEASROGAEYDMRHIFT-----GTLVEADPFHATVILVANHDTIQLQALE 340  
Qy 295 -----WNKYPAYAFILTYE-QOPTIFYED-----YEELNK-----DKLNLIW 332  
Db 341 APVEPWFKPLAYALILLRENGVPSVFYDLYGASYEDNGENGETCRVDMPEVINGOLDRLIL 400  
Qy 333 IHENLAGGSTDIVYDNDLI-FVRNGYDGPGLITVINGSSKAGRWVYVVPFAGACIH 391  
Db 401 ARQRFAGH-IQTLFFDHPNCIAFSRSGTEENPGCVVLSNGDDGDKETLLGLDNYANKTW 459  
Qy 392 EYTG 408  
Db 460 DFLGN 472  
RESULT 4  
A24549  
alpha-amylase (EC 3.2.1.1) precursor - Bacillus stearothermophilus (strain NZ-3)  
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C:Species: Bacillus stearothermophilus





Wed Jun 30 10:27:03 2004

C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: B85810  
 R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: B85810  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-495 <STO>  
 A:Cross-references: GB:AE005174; NID:gl2516028; PIDN:AAG56942.1; GSPDB:GN00145; UWGP:Z30  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: amyA  
 C:Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology

Query Match 18.1%; Score 444.5; DB 2; Length 495;  
 Best Local Similarity 28.7%; Pred. No. 1.3e-24;  
 Matches 131; Conservative 67; Mismatches 140; Indels 119; Gaps 21;

QY 13 IMQAFYWDVPSGGIMWDTIRQIPKIPWYDAGISAIWIPKASKMGGMGYSGYDPYDFDLG 72  
 DB 6 LLOCFHWYYPGEGKLPWELAEADGFNDIGINWMLPPAYKASGGYSGYDSDYDLFDLG 65  
 QY 73 EYDQKCTVTRFGSKOELVNMINTAHAYGMKVIADIVINHRAGGDL-----WNEFFVNDY 127  
 DB 66 EFDQKGSIPKTYGDKVQLLAADALKRNDIAVLDDVNNHKGADKEAIRVQRVNADDR 125  
 QY 128 T-----WTDFSKVA-SGKYTANYLDF-----HPNELHAGDSGTF----- 160  
 DB 126 TQDEEIIIECEGRTYTFPARAGQNSQFWDKCFSGIDHIENPNE-----DGIFKIYND 180  
 QY 161 ---GGYPDICHDK--SWD-----QYWL-WASQESYAAYLRSIGIDAM 196  
 DB 181 YTCEGNNDQVDDDELGNFDYLMGENIDFRNHAVTEEEKYWARWVMEQTQC-----DGF 232  
 QY 197 RFDYVKGYPWVVKDMLNHWGWA-----VGEYMDTNVDALNWAYSSGAK--VFDFAL 248  
 DB 233 RLDAVKHHPAWFYKWEIHEVQEVAPKPLFIVAEYWSHEVDKQTYIDQVEGKTNLFDAPL 292  
 QY 249 YIKWDEA-----FONKNIPALVSLQNGQTVVSRDPFKAVTFVANHDTDI----- 294  
 DB 293 QMKFHEASRGRNYD-----MTQIFTG-TLVEADXFHAVTXVANHDTQXLALEAPVE 344  
 QY 295 -WNKYPAYAFILTYE-GQPTIFYRD-----YEELWNK-----DKLNLIWIHEN 336  
 DB 345 PWFKPLAYALLIRENGVPSVFPYDLYGAHYEDVGGDGTYPIDMPIIEQLDELILARQ 404  
 QY 337 LAGGSTDIVYDNDLFI-FVRNGYGDKPGLLITYINLG 372  
 DB 405 FAHG-VOTLFFPHNCIAFSRGTDEVPCCGVVMSNG 440

Search completed: June 29, 2004, 09:36:16  
 Job time : 23 secs

Db 295 SFHDASKQGGDFDMRSIFDGLSLVAV-----PDHAVTLVDNHDTPQLQSLEAPVE 344  
 QY 295 -WNKYPAYAFI-LTYEGQTIYRD-----YEWLNK-----DKLKNLIWIHEN 336  
 Db 345 PWFKPLAYAILLREEGVPCVFPDLFGTSTYDTGNDGNEYKIDIPAEICLPKLEARSR 404  
 QY 337 LAGGSTDIVYDNDDELIFVRNGYGDKEPLITYINLGSSKAGRWVYVPKFGACIHEYTGN 396  
 Db 405 FANGPQTDIFDDASCIARHGTADAPGCVVMSNGEPGEKQADLGPERRAGSVWRDFLGH 464

RESULT 13  
 G98247  
 cytoplasmic alpha-amylase (1,4-alpha-D-glucan glucanohydrolase) [imported] - Agrobacterium  
 C/Species: Agrobacterium tumefaciens  
 C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
 C/Accession: G98247  
 R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
 A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A/Reference number: A97359; MUID:21608551; PMID:11743194  
 A/Accession: G98247  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-506 <KUR>  
 A/Cross-references: GB:AE007870; PIDN:AAK89505.1; PID:g15159379; GSPDB:GN00170  
 C/Genetics:  
 A/Gene: AGR\_L1863  
 A/Map position: linear chromosome  
 C/Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology

Query Match 18.4%; Score 451.5; DB 2; Length 506;  
 Best Local Similarity 27.3%; Pred. No. 4.1e-25;  
 Matches 131; Conservative 70; Mismatches 168; Indels 111; Gaps 15;

QY 10 GGVMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWTPPASKMGAGYSGYDYPDF 69  
 Db 14 GRLLLOFFHHYYPDGGKLVSEVAKESLAKGIVDVLPPAYKGAAGYSYGYDTYDLF 73  
 QY 70 DLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVNHRAGGDLNPFV----- 124  
 Db 74 DLGEYDQKGTATKYGDRAALSHAGKTLKNDGIRVTHVNLNKGGADEKEXVRVRNP 133  
 QY 125 ---NDYTWTFDFSKVSGKKT-----ANYL 145  
 Db 134 DRTDIDDEFPALATRTFFPGRNGHSHKFTWDLKCFSGVDHIEPTEDGIFRLNBYG 193  
 QY 146 DPHNELHAGDSGTF----GGYDICHDKSDQ--YW-LWASQESYAAYLRSIGIDAWRF 198  
 Db 194 DGEWNEEVDOENGDFYLMGADVEFPNRVAYEELKXGWRWLSQ-----VQVDGFL 245  
 QY 199 DIVKGYAPVWVKDLNWWCG-----WAVGETWDTNVDVNLW--AYSSGAKVDFALYY 250  
 Db 246 DAAKHIPAWFFRDVGHMRETVDPLDFVVAEYWHPDLEALKSYLBDVKQLMLFDVALHH 305  
 QY 251 KMDEA-----FDKNKI--PALVSALQNGQTVVSRDPPKAVTFVANHDTDII----- 294  
 Db 306 SFHDASKQGGDFDMRSIFDGLSLVAV-----PDHAVTLVDNHDTPQLQSLEAPVE 355  
 QY 295 -WNKYPAYAFI-LTYEGQTIYRD-----YEWLNK-----DKLKNLIWIHEN 336  
 Db 356 PWFKPLAYAILLREEGVPCVFPDLFGTSTYDTGNDGNEYKIDIPAEICLPKLEARSR 415  
 QY 337 LAGGSTDIVYDNDDELIFVRNGYGDKEPLITYINLGSSKAGRWVYVPKFGACIHEYTGN 396  
 Db 416 FANGPQTDIFDDASCIARHGTADAPGCVVMSNGEPGEKQADLGPERRAGSVWRDFLGH 475

RESULT 14  
 A45738  
 alpha-amylase (EC 3.2.1.1), cytosolic - Escherichia coli (strain K-12)  
 N/Alternate names: 1,4-alpha-D-glucan glucanohydrolase

C/Species: Escherichia coli  
 C/Date: 07-Apr-1994 #sequence\_revision 31-Oct-1997 #text\_change 01-Mar-2002  
 C/Accession: D64956; A45738  
 R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
 A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A/Title: The complete genome sequence of Escherichia coli K-12.  
 A/Reference number: A64720; MUID:97426617; PMID:9278503  
 A/Accession: D64956  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-495 <BLAT>  
 A/Cross-references: GB:AE000285; GB:U00096; NID:g1788229; PIDN:AACT4994.1; PID:g1788236;  
 A/Experimental source: strain K-12, substrain M61655  
 R/Raba, M.; Kawagishi, I.; Mueller, V.; Kihara, M.; Macnab, R.M.  
 J. Bacteriol. 174, 6644-6652, 1992  
 A/Title: Escherichia coli produces a cytoplasmic alpha-amylase, amyA.  
 A/Reference number: A45738; MUID:93015717; PMID:1400215  
 A/Accession: A45738  
 A/Molecule type: DNA  
 A/Residues: 1-18,'SS',21-108,'V',110-148,'E',150-233,'I',235-495 <RAH>  
 A/Cross-references: GB:L01642; NID:g146021; PIDN:AAA23810.1; PID:g146023  
 C/Genetics:  
 A/Gene: amyA  
 C/Function:  
 A/Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
 A/Pathway: glycogen/starch degradation  
 C/Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
 C/Keywords: cytosol; glycosidase; hydrolase; polysaccharide degradation  
 F/202-335/domain: alpha-amylase core homology <AMY>  
 F/104,204,239/Binding site: calcium (Asn, Asn, His) #status predicted  
 F/235,265,332/Active site: Asp, Glu, Asp #status predicted

Query Match 18.2%; Score 447.5; DB 1; Length 495;  
 Best Local Similarity 28.7%; Pred. No. 7.8e-25;  
 Matches 131; Conservative 68; Mismatches 139; Indels 119; Gaps 21;

QY 13 IMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWTPPASKMGAGYSGYDYPDFDLG 72  
 Db 6 LLQCFHHYYPGGKLVPELAERADGNDIGINWVLPAYKAGSGYSGYDYPDFDLG 65  
 QY 73 BYDQKGVETRFSGKQELVNMINTAHAYGMKVADIIVNHRAGGDLN-----WNPFVNDY 127  
 Db 66 EFDQKGSIPKYGDKAQLLAIDALKRNDIAVLDDVVVNHKMGGADEKAIQRQVRNADDR 125  
 QY 128 T-----WTDFSKVA-SGKTYANYLDF-----HPNELHAGDSGTF----- 160  
 Db 126 TDDEEIECEGWTRYTFPPARAGYSQFIWDFKFCSDIHENPDE-----DGIFKIVND 180  
 QY 161 ---GGYDICHDK-SWD-----OYWL-WASQESYAAYLRSIGIDAW 196  
 Db 181 YTGEGWNDQVDDBLGNFDYLMGENIDFRNHAVTEEEKYARWYMEQTQC-----DGF 232  
 QY 197 RFDYVKGYPVWVKDLNWWGWA-----VGEYWDNTVDVNLWAYSSGAK--VFDFAL 248  
 Db 233 RLDAVKHIIPAWFKIEWTEHVEVAPKPLFIVAYNSHEVDKLTQYTDQVEGKTMFLDAPL 292  
 QY 249 YKKDEA-----FDKNIPALVSALQNGQTVVSRDPPKAVTFVANHDTDII----- 294  
 Db 293 QMKFHEASRMGRDYD-----MTQIFTG-TLVEADPFHATVILVANHDTQPLQALEAPVE 344  
 QY 295 -WNKYPAYAFIYE-GQPTIFYRD-----YEWLNK-----DKLKNLIWIHEN 336  
 Db 345 PWFKPLAYAILLRENGVSVFYPDLIGHAYHVEDGGGQTYPIDMPIISOLDLILARQR 404  
 QY 337 LAGGSTDIVYDNDDELIFVRNGYGDKEPLITYINLG 372  
 Db 405 FAHGG-VQTLFFDHPNCIAFSRSGTDFPFGCVVMSNG 440

RESULT 15  
 B85810  
 cytoplasmic alpha-amylase [imported] - Escherichia coli (strain O157:H7, substrain EDL93)

QY 11 GVIMQAFVMDVPSGGIMWDTIRQKIPWYDAGISAIWIPASKGMGGAYSMGYDPYDFD 70  
 DB 40 GTMGOYFEWYLPNDGNHNLNSDASNLKSGITAVWIPPAWKG-ASQNDVGYGAYDLYD 98  
 QY 71 LGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGD-----LEWNP- 122  
 DB 99 LGEFNQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGD-----LEWNP- 158  
 QY 123 -----FUNDYT---WTFPSKVASCK-----YTANYLDHPHNLHAGDSTGEGYD 165  
 DB 159 NKNQEVTEYIEATWTRFPCGRNTHSPKRWYHFDGVDWQSRRLNRYKFRG----- 215  
 QY 166 ICHDKSWD-----QYWLNAS-----QESYAAIYRSIGIDAWRFYVKG 203  
 DB 216 --HGKANDWEVDTENGNYDYLMDYADIMDHPEVNNELRNWGVYNTLTGLDGRIDAVKH 273  
 QY 204 YAPVWVKDNLNMGW-----WAVEGYMDTNVDAV-----LNWYSSGAKVDFALYK 251  
 DB 274 IKYGFTRDWINHRSATGKMFVAEFKNDLGALENYLOKTNWHS-----VFDVPLHYN 329  
 QY 252 MDEA-----FDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTD-----IWN 296  
 DB 330 LYNASKSGGNDYMRNI-----FNG-TVVRPHSHAVTFVDNHDSPPEALESFVEWF 381  
 QY 297 KYPAYAFILTYE-COPTIFYRDIYEW-----LNKOKLXLIWIHENLAGGSTDIVYDN 349  
 DB 392 KPLAYALTREQQSPSVFYDGYGIPTHGVPMARSKIDILEARQYAYGKON-DYLDH 440  
 QY 350 DELI-FVRNGYGDKP--GLITYINLGSSKAGRWYVYPK-PAGACIHEYTGNLGGWVKTV 405  
 DB 441 HNIIGWTRGNTAHENSGLATINSKGAG-GSKWMEVGRNKGAGQVSDITGNRTGTW--I 497  
 QY 406 YSSGW 410  
 DB 498 NADGW 502  
 RESULT 11  
 B90962  
 Cyttoplasmic alpha-amylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C:Accession: B90962  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: B90962  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-495 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA036089.1; PID:g13362134; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:  
 C:Superfamily: alpha-amylase, amylioliquefaciens type; alpha-amylase core homology  
 A:Gene: ECs2666  
 Query Match 18.4%; Score 452.5; DB 2; Length 495;  
 Best Local Similarity 28.9%; Pred. No. 3.4e-25;  
 Matches 132; Conservative 67; Mismatches 139; Indels 119; Gaps 21;  
 QY 13 IMQAFVMDVPSGGIMWDTIRQKIPWYDAGISAIWIPASKGMGGAYSMGYDPYDFD 72  
 DB 6 LQCFHMYEPEGKMLPELAERADGNDIGINWMLPPAYKAGSGSYGVYDYLDF 65  
 QY 73 EYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGD-----WNPFDVND 127  
 DB 66 EFDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGD-----WNPFDVND 125  
 QY 128 T-----WTFPSKVASCK-----YTANYLDHPHNLHAGDSTGEGYD 160  
 DB 126 TDDEEIEEGWTRITFPARQYSGFQWDFKCFSGDIHNPNE-----DGIFKIVND 180

QY 161 ---GGYDICHDK--SWD-----QYWL-WASQESYAAIYRSIGIDAW 196  
 DB 181 YTBGMNDQVDDSLGNPDYLMGENIDFRNHAVTEETKYWARWYMEQTC-----DGF 232  
 QY 197 RFDYVKGAYFVWVKDNLNMGW-----VGEYMDTNVDAVLNWAYSSGAK--VPDFAL 248  
 DB 233 RLDAVKHIPAWFKWETEHVQEVAPKPLFVAEYMSHEVDKLQTYIDQVEGKTMFLDAPL 292  
 QY 249 YKMDER-----FDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDI 294  
 DB 293 QMKPHEASRMGRNY-----MTQIFTG-TLVEADPFHAVTLVANHDTQPLQALEAPVE 344  
 QY 295 -WKKYPAYAFILTYE-COPTIFYRDIYEW-----LNKOKLXLIWIHEN 336  
 DB 345 PWFKPLAYALILLRENGVPSVFPDLYGAHYEDVGGDGTQYFIDMPTIEQLDELILARQ 404  
 QY 337 LAGSTDIVYDNDELI-FVRNGYGDKPGLITYINLG 372  
 DB 405 FAHG-VQTLFFDHPNCIAFSRSGTDEYPCGVVMSNG 440  
 RESULT 12  
 AD3038  
 alpha-amylase amyA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 19-Nov-2002  
 C:Accession: AD3038  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.  
 : Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan  
 : Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AD3038  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-495 <KUR>  
 A:Cross-references: GB:AB008689; PIDN:AAL44722.1; PID:g17742354; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: amyA  
 A:Map position: linear chromosome  
 C:Superfamily: alpha-amylase, amylioliquefaciens type; alpha-amylase core homology  
 Query Match 18.4%; Score 451.5; DB 2; Length 495;  
 Best Local Similarity 27.3%; Pred. No. 4e-25;  
 Matches 131; Conservative 70; Mismatches 168; Indels 111; Gaps 15;  
 QY 10 GGVMQAFVMDVPSGGIMWDTIRQKIPWYDAGISAIWIPASKGMGGAYSMGYDPYDF 69  
 DB 3 GRTLLQFHHYYPDGGKLVSEVAEKAESLAKWGITDVLPPAYKAGSGSYGVYDYLDF 62  
 QY 70 DICEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEWNPV 124  
 DB 63 DLGEFDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEWNPV 122  
 QY 125 ---NDYTWTFDPSKVASCKY-----YTANYLDHPHNLHAGDSTGEGYD 145  
 DB 123 DDTDTIDDEDFPALAYTRTFPGRNGKSKFTINDLCKFCGVHDHIEPTEDGIFRLVNEYG 182  
 QY 146 DFPNELHAGDSGTF-----GGYPDICHKSWDQ--YW-LWASQESYAAIYRSIGIDAWRF 198  
 DB 183 DGSWNEVDQNGNFDYLMGADVSRNRAVYBELAYKRWLSEQ-----VQVDGFL 234  
 QY 199 DYVKGAYFVWVKDNLNMGW-----WAVEGYMDTNVDAVLN-----AYSSGAKVDFALY 250  
 DB 235 DAAKHIPANFRDWHGHRBTVDPLFVVAEYHHPDLSALKSYLELVKQLMLFDVALHH 294  
 QY 251 KMDER-----FDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDI 294

### Query Match

[illegible]

## RESULT 8

ALSPF  
alpha-amylase (EC 3.2.1.1) precursor - *Bacillus stearothermophilus* (strain DY-5) plasmid  
N;Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C;Species: *Bacillus stearothermophilus*  
C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Feb-1997  
C;Accession: A91999; B91999; A91804; A00845  
R;Ihara, H.; Sasaki, T.; Tsuboi, A.; Yamagata, H.; Tsukagoshi, N.; Uda, S.  
J. Biochem. 98, 95-103, 1985  
A;Title: Complete nucleotide sequence of a thermophilic alpha-amylase gene: homology bet  
A;Reference number: A91999; MUID:86008166; PMID:3876333  
A;Accession: A91999  
A;Molecule type: DNA  
A;Residues: 1-548 <IH1>  
A;Cross-references: GB:X02769  
A;Experimental source: plasmid pHI300 from strain DY-5  
A;Accession: B91999  
A;Molecule type: protein  
A;Residues: 35-48 <IH2>  
A;Experimental source: strain DY-5  
R;Tsukagoshi, N.; Iritani, S.; Sasaki, T.; Takemura, T.; Ihara, H.; Idota, Y.; Yamagata, J. *Bacteriol.* 164, 1182-1187, 1985  
A;Title: Efficient synthesis and secretion of a thermophilic alpha-amylase by protein-p  
A;Reference number: A91804; MUID:86059211; PMID:2999073  
A;Contents: pBAM101  
A;Accession: A91804  
A;Molecule type: DNA  
A;Residues: 1-29, 'Q', 31-75, 'W', 77-122 <TSU>  
C;Comment: Alpha-amylase genes have been found on plasmids and in multiple copies on the  
C;Genetics:  
A;Genome: plasmid  
A;Start codon: GTG  
C;Function:  
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A;Pathway: glycogen/starch degradation  
C;Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
C;Keywords: extracellular protein; glucosidase; heat-stable protein; hydrolase; polysac

F1-1-34/Domain: signal sequence #status predicted <SIG>  
F135-348/Product: alpha-amylase #status experimental <MAT>  
F235-368/Product: alpha-amylase core homology <AMY>  
F139-237, 272/Binding site: Calcium (Asp, Asp, His) #status predicted  
F1268, 298, 345/Active site: Asp, Glu, Asp #status predicted

Query Match 18.7%; Score 459; DB 1; Length 548;  
Best Local Similarity 28.6%; Pred. No. 1.3e-25;  
Matches 148; Conservative 67; Mismatches 162; Indels 140; Gaps 26;

[illegible]

## RESULT 9

alpha-amylase [EC 3.2.1.1] precursor [validated] - Bacillus licheniformis  
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C:Species: Bacillus licheniformis  
C:Date: 30-Jun-1987  
C:Sequence revision 24-Apr-1998 #text-change 15-Sep-2000  
C:Accession: A91997; B24549; A91796; A21663; I39772; A26151; S53768; A00844  
R:Yukio, T.; Nomura, T.; Tezuka, H.; Tsuboi, A.; Yamagata, H.; Tsukagoshi, N.; Udaoka, S.  
J. Biochem. 98, 1147-1156, 1985  
A:Title: Complete nucleotide sequence of a gene coding for heat- and pH-stable alpha-amylase deduced from the DNA sequences.  
A:Reference number: A91997; MUID:86111694; PMID:2418011  
A:Accession: A91997  
A:Molecule type: DNA  
A:Residues: 1-162, 'R', 164-512 <YUU>  
A:Cross-references: GB:I30236; NID:g39551; PIDN:CAA26981.1; PID:g39552  
R:Gray, G.L.; Mainzer, S.E.; Revy, M.W.; Lamsa, M.H.; Kindie, K.L.; Carmona, C.; Requaadt, J. Bacteriol. 166, 635-643, 1986  
A:Title: Structural genes encoding the thermophilic alpha-amylases of Bacillus stearothermophilus  
A:Reference number: A91817; MUID:86195857; PMID:3009417  
A:Accession: B24549  
A:Molecule type: DNA  
A:Residues: 1-338, 'G', 340-348, 'S', 350-512 <GRA>  
A:Cross-references: GB:M13256; NID:g142510; PIDN:AAA22240.1; PID:g142511  
A:Experimental source: NCIB 8061

QY 123 -----FVNDY---TWTFDSKVASGKYTANYLDHPHNEH----- 153  
Db 153 NRQETSEEIOIKAWIDFPFGRGN---TYSDFKWHYHFDGADWDESKISRIKFRGE 209  
QY 154 -----AGDSCTFG---GYPDICHD-----KSWDQYWLWASQESYAAVLRISGIDA 195  
Db 210 GKAWDEWVSSENGNYDYLWADVDYDHPDVVAETKKWG---IW-----YANLSLDG 258  
QY 196 WRFDYVGYAPVWVKDWL-----NWMGWAGVEYDWDVAVLW---AYSGAKVFPFA 247  
Db 259 FRIDAAGKHIFSLRDWQVQVQATCKEMFTVAEYQNNAGLENYLNKTSFNQGVFDVP 318  
QY 248 LYYKMDA-----FNNKNIIPALVSALQNGQTVVSRDPFKAVTFVANHDTD----- 292  
Db 319 LHENLQAASSGGGYDMR-----LLDG-TVSSHPEKATFVENHDTQCGQLESTV 370  
QY 293 IWNKYPAYAFILTYE-GOPTIPYRDY-----EWNLNKDKLNLIWIHENLAGGST 342  
Db 371 QTWFKELAYAFILTYREGSPQVFGYMGYKGTSPKEIPSLKDNIEPIKARKEYAYGPQ 430  
QY 343 DIVVYDNDLI-FVRNGYGD--KPGGLITYINLGSSKAGWVVPKFPAGACIHEYTGNLGG 399  
Db 431 H-DYIDHPDVIGWTRGDSAAKSGLAALITDQPGSKWYAGLNAGETWYDITGNRSD 489  
QY 400 WDKYVYSSGW 410  
Db 490 TVK--IGSDGW 498

RESULT 6  
A:Alpha-amylose (EC 3.2.1.1) precursor - Bacillus stearothermophilus (strain DNI1792)  
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C:Species: Bacillus stearothermophilus  
C:Date: 28-Oct-1994 #sequence\_revision 18-Aug-1995 #text\_change 13-Jun-1997  
C:Accession: A54541  
R:Jorgensen, P.L.; Poulsen, G.B.; Diderichsen, B.  
FEMS Microbiol. Lett. 77, 271-276, 1991  
A:Title: Cloning of a chromosomal alpha-amylose gene from Bacillus stearothermophilus.  
A:Reference number: A54541  
A:Accession: A54541  
A:Molecule type: DNA  
A:Residues: 1-549 <OR>  
A:Cross-references: GB:X59476  
A:Experimental source: chromosomal DNA of strain DNI1792  
C:Comment: Alpha-amylose genes have been found on plasmids and in multiple copies on the  
C:Genetics:  
A:Start codon: GTG  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
C:Superfamily: alpha-amylose, amyloliquefaciens type; alpha-amylose core homology  
C:Keywords: extracellular protein; glycosidase; heat-stable protein; hydrolase; polysac  
F:1-34/Domain: signal sequence #status predicted <SIG>  
F:35-549/Product: alpha-amylose #status predicted <MAT>  
F:235-368/Domain: alpha-amylose core homology <AMY>  
F:139,237,272/Binding site: calcium (Asp, Asp, His) #status predicted  
F:268,298,365/Active site: Asp, Glu, Asp #status predicted

Query Match 19.0%; Score 466.5; DB 1; Length 549;  
Best Local Similarity 28.4%; Pred. No. 3.7e-26;  
Matches 146; Conservative 67; Mismatches 168; Indels 133; Gaps 25;

QY 11 GVIMQAFYDVPDGGIWDITRQKIPENYDAGISAIWIPPPASKMGAGYSGYDYPDFD 70  
Db 40 GFMQGYFYLPPDDGTLTKWVANEANNLSLIGITALWLPYAKGTSRS-DVGYGYDLYD 98  
QY 71 LGEYDQKGTVEPRFGSKOELNMINHTAHAYGMKVIADIVINHRAGGD-LEW-----NPF 123  
Db 99 LGEFNQKGVTRKYTKAQYLQAIQAAHAGQVYADVVFHKGADGTWVDAVENPS 158  
QY 124 VND-----YTWTFDSKVASGKYTANYLDHPHNEHAG----- 155

Db 159 DRNOEISGTVOIQAWTKDFPGRGN---TYSSFKRWYHFDGVDWDESKLSRIYKFRGI 215  
QY 156 -----DSGTFG-----GYPDICHD--KSWDQYWLWASQESYAAVLRISGIDA 195  
Db 216 GKAWDEWVDENGNYDYLWADLMDHPVVVTELKNGKW-----YVNTNIDG 264  
QY 196 WRFDYVGYAPVWVKDWLWGG-----WAVEYDWDVAVLW---AYSGAKVFPFA 247  
Db 265 FRIDAAGKHIFSPFPDWSYVRSQTGKPLFTVGEYNSYDINKLHNYITKTGDTMSLFDAP 324  
QY 248 LYYKMDA-----DEAFDNKNIIPALVSALQNGQTVVSRDPFKAVTFVANHDTD----- 292  
Db 325 LHNFYTASKSGGAFDNRILTMT-----NTLMKQDPTLAVTFVDNHDIEPQALOSWV 376  
QY 293 IWNKYPAYAFILTYE-GOPTIPYRDY---EWNLNKDKLNLIWIHENLAGGSTDI 345  
Db 377 DPWFKELAYAFILTYREGSPQVFGYMGYKGTSPKEIPSLKDNIEPIKARKEYAYGPQ 435  
QY 346 YVDNDELII-FVRNGYGD--GLITYINLGSSKAGWVVPKFPAGACIHEYTGNLGGW 401  
Db 436 YLHSDIIGWTRGEGTKPGSGLAALITDQPG--GSKMWYVQKQAGKVFYDITGNRSDTV 494  
QY 402 DKYVYSSGWYLEAPAYDPANGQY-----GYSVW 430  
Db 495 T--INSDGW-----GDFKNGSGSVSW 514

RESULT 7  
A:Alpha-amylose (EC 3.2.1.1) precursor - Bacillus stearothermophilus plasmid pATs  
N:Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
C:Species: Bacillus stearothermophilus  
C:Date: 05-Jun-1987 #sequence\_revision 18-Aug-1995 #text\_change 18-Jun-1999  
C:Accession: A24436; I39777  
R:Nakajima, R.; Imanaka, T.; Aiba, S.  
J. Bacteriol. 163, 401-406, 1985  
A:Reference number: A24436; PMID:85234394; PMID:3924897  
A:Accession: A24436  
A:Molecule type: DNA  
A:Residues: 1-549 <NAK>  
A:Cross-references: GB:M11450  
A:Experimental source: plasmid pATs  
A:Note: amino end of the mature protein also determined  
R:Jorgensen, P.L.; Hansen, C.K.; Poulsen, G.B.; Diderichsen, B.  
Gene 96, 37-41, 1990  
A:Title: In vivo genetic engineering: homologous recombination as a tool for plasmid con  
A:Reference number: I39772; PMID:9102499; PMID:2265757  
A:Accession: I39777  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-45 <RES>  
A:Cross-references: GB:M62638; NID:GI42514; PIDN:AAA2242.1; PID:GI42515  
C:Comment: Alpha-amylose genes have been found on plasmids and in multiple copies on the  
C:Genetics:  
A:Gene: amyS  
A:Genome: plasmid  
A:Start codon: GTG  
C:Function:  
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A:Pathway: glycogen/starch degradation  
C:Superfamily: alpha-amylose, amyloliquefaciens type; alpha-amylose core homology  
C:Keywords: extracellular protein; glycosidase; heat-stable protein; hydrolase; polysac  
F:1-34/Domain: signal sequence #status predicted <SIG>  
F:35-549/Product: alpha-amylose #status experimental <MAT>  
F:235-368/Domain: alpha-amylose core homology <AMY>  
F:139,237,272/Binding site: calcium (Asp, Asp, His) #status predicted  
F:268,298,365/Active site: Asp, Glu, Asp #status predicted

Query Match 18.8%; Score 461.5; DB 1; Length 549;  
Best Local Similarity 28.4%; Pred. No. 8.6e-26;  
Matches 146; Conservative 65; Mismatches 170; Indels 133; Gaps 25;

QY 11 GVIMQAFYDVPDGGIWDITRQKIPENYDAGISAIWIPPPASKMGAGYSGYDYPDFD 70



C;Date: 30-Jun-1998 #sequence revision 18-Aug-1995 #text\_change 18-Jun-1999  
 C;Accession: A24549; I39501; I39770  
 R;Gray, G.L.; Mainzer, S.E.; Rey, M.W.; Lamsa, M.H.; Kindie, K.L.; Carmona, C.; Requadt,  
 J. Bacteriol. 166, 635-643, 1986  
 A;Title: Structural genes encoding the thermophilic alpha-amylases of Bacillus stearother-  
 A;Reference number: A31817; MUID:86195857; PMID:3009417  
 A;Accession: A24549  
 A;Molecule type: DNA  
 A;Residues: 1-549 <GRA>  
 A;Cross-references: GB:M13255; NID:g142512; PIDN:AAA22241.1; PID:g142513  
 A;Experimental source: Genomic DNA of strain NZ-3  
 R;Satch, H.; Nishida, H.; Isono, K.  
 J. Bacteriol. 170, 1034-1040, 1988  
 A;Title: Evidence for movement of the alpha-amylase gene into two phylogenetically dista  
 A;Reference number: I39501; MUID:86139156; PMID:3257753  
 A;Accession: I39501  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 536-549 <RES>  
 A;Cross-references: GB:M29577; NID:g142476; PIDN:AAA22225.1; PID:g142478  
 A;Experimental source: strain DY-5  
 A;Accession: I39770  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 536-549 <RE2>  
 A;Cross-references: GB:M29578; NID:g142484; PIDN:AAA22228.1; PID:g142486  
 A;Experimental source: strain 799  
 C;Comment: Alpha-amylase genes have been found on plasmids and in multiple copies on the  
 C;Genetics:  
 A;Start codon: GTG  
 C;Function:  
 A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
 A;Pathway: glycogen/starch degradation  
 C;Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
 C;Keywords: extracellular protein; glycosidase; heat-stable protein; hydrolase; polysac  
 F;1-34/DNA: signal sequence #status predicted <SIG>  
 F;35-549/Product: alpha-amylase #status predicted <MAT>  
 F;235-368/DNA: alpha-amylase core homology <AMY>  
 F;139,237,272/Binding site: calcium (Asp, Asp, His) #status predicted  
 F;268,298,365/Active site: Asp, Glu, Asp #status predicted  
 Query Match 19.3%; Score 475.5; DB 1; Length 549;  
 Best Local Similarity 29.2%; Pred. No. 8.3e-27;  
 Matches 151; Conservative 61; Mismatches 166; Indels 139; Gaps 25;  
 QY 11 GVIMQAFYHVPSSGIGTIRQKIPWYDAGISAIWIPPPSKMGGMGAYSGYDYPDFD 70  
 DB 40 GTMMQYFEWYLPDDGTLTKVANEANNLSSGLTALWLPAYKGRS-DVGYGYDLYD 98  
 QY 71 LGEYDQKGVETFRFGSKQELVNMINTAHAYGMKVADIVINHRRAGD-LEW-----NPF 123  
 DB 99 LGEFNQKGVIRKYGTQKAYQALQIAAAGAGQVAVVDFDKGAGDGTWVDVAVENP- 157  
 QY 124 VNDYTDPSKVASGYTANYLDHFENELHAGSGTFGYPDICHKSNQY----- 175  
 DB 158 -----SDRNOETSGTY-----QIQAATKDFNGRNGTYSFKRWVWYHFGVDWD 201  
 QY 176 -----WLWASQES-----YAA-----YLRSIG 192  
 DB 202 ESKLSRIYKFRIGKANDWEDVTENGNYDLYMAYDLMDHPEVVTTELKMGKWIYNTN 261  
 QY 193 IDAWREDYKGYAPVWVVKWLNWGG-----WAVEGYWDTWDAVLNWAYSSGA--KVF 244  
 DB 262 IDGFRDLAVKHIFKSPFPDLVSYSQTKPLFTVGEYSYDINKLHNYITKNGTMSLF 321  
 QY 245 DFALYKM-----DEAFDNKIPALVSAALQNGQTVSRDRPKAVTFVANHTD----- 292  
 DB 322 DAPLHNKFTYASKSGGAF-----MSTLMN-NTLUNKOQPLAVTVNHDHTEGQALQ 373  
 QY 293 ---IWNKYPAVAFILT-YEQPTIFYRDE--EWLN-----KDKLKNLIWIHENTJAGST 342  
 DB 374 SWDPWFKPLAYAFILTRQSGYPCVFGVDYGYGPQYNIPSLKSKIDPILLIARRDYAYG-T 432

QY 343 DIVYDNDLI-FVRNGYGDKP--GLITVINLGSSKAGRWVYVPR-FAGACIHEYTGNLG 398  
 DB 433 QHDYLDHSDIIGTWREGVTEKPGSLAALITDGP-GSKWMTVGKHAGKVFYDLTGNS 491  
 QY 399 GWYDKYVSSGVWVYLEAPAYDPANQY-----GYSVW 430  
 DB 492 DTVT--INSDGW-----GFEKVGSGSVW 514  
 RESULT 5  
 ALSN  
 alpha-amylase (EC 3.2.1.1) precursor - Bacillus amyloliquefaciens  
 N;Alternate names: 1,4-alpha-D-glucan glucanohydrolase  
 C;Species: Bacillus amyloliquefaciens  
 C;Date: 30-Nov-1980 #sequence revision 30-Jun-1987 #text\_change 18-Jun-1999  
 C;Accession: A93889; A93037; I39756; I39763; A00843  
 R;Takkinen, K.; Pettersson, R.F.; Kalkkinen, N.; Palva, I.; Soderlund, H.; Kaariainen, L.  
 J. Biol. Chem. 258, 1007-1013, 1983  
 A;Title: Amino acid sequence of alpha-amylase from Bacillus amyloliquefaciens deduced fr  
 A;Reference number: A93889; MUID:83108808; PMID:6185474  
 A;Contents: pUB110  
 A;Accession: A93889  
 A;Molecule type: DNA  
 A;Residues: 1-514 <TAK>  
 A;Cross-references: GB:J01542; GB:J01543; GB:M12033; GB:M12034; NID:g142428; PIDN:AAA2215  
 R;Chung, H.S.; Friedberg, F.  
 Biochem. J. 185, 387-395, 1980  
 A;Title: Sequence of the N-terminal half of Bacillus amyloliquefaciens alpha-amylase.  
 A;Reference number: A90307; MUID:80241725; PMID:6156671  
 A;Accession: A90307  
 A;Molecule type: protein  
 A;Residues: 32-53, 1, 55-63, 1, 65-78, 1, 80-83, 1, 85-222 <CHU>  
 R;Palva, I.; Pettersson, R.F.; Kalkkinen, N.; Lehtovaara, P.; Sarvas, M.; Soderlund, H.;  
 Gene 15, 43-51, 1981  
 A;Title: Nucleotide sequence of the promoter and NH2-terminal signal peptide region of t  
 A;Reference number: I39756; MUID:82051296; PMID:6170539  
 A;Accession: I39756  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-96 <RES>  
 A;Cross-references: EMBL:V00092; NID:g39297; PIDN:CAA23430.1; PID:g39298  
 R;Ruhonen, L.; Hackman, P.; Lehtovaara, P.; Knowles, J.K.C.; Karaenen, S.  
 Gene 59, 161-170, 1987  
 A;Title: Efficient secretion of Bacillus amyloliquefaciens alpha-amylase cells by its ow  
 A;Reference number: I39763; MUID:88137952; PMID:2830166  
 A;Accession: I39763  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-39 <RE2>  
 A;Cross-references: GB:M18424; NID:g142430; PIDN:AAA22192.1; PID:g142431  
 C;Function:  
 A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
 A;Pathway: glycogen/starch degradation  
 C;Superfamily: alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  
 C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation  
 F;1-31/DNA: signal sequence #status predicted <SIG>  
 F;32-514/Product: alpha-amylase #status predicted <MPT>  
 F;229-362/DNA: alpha-amylase core homology <AMY>  
 F;133,231,266/Binding site: calcium (Asn, Asp, His) #status predicted  
 F;262,292,359/Active site: Asp, Glu, Asp #status predicted  
 Query Match 19.2%; Score 472.5; DB 1; Length 514;  
 Best Local Similarity 29.9%; Pred. No. 1.3e-26;  
 Matches 147; Conservative 60; Mismatches 157; Indels 117; Gaps 22;  
 QY 11 GVIMQAFYHVPSSGIGTIRQKIPWYDAGISAIWIPPPSKMGGMGAYSGYDYPDFD 70  
 DB 34 GTLMQYFEWYLPDDGTLTKVANEANNLSSGLTALWLPAYKGRS-DNGYGYDLYD 92  
 QY 71 LGEYDQKGVETFRFGSKQELVNMINTAHAYGMKVADIVINHRRAGD-----LEWNP- 122  
 DB 93 LGEFOOKGVTRTKYKSELODAIGLSHRNVQYGVNLNKKACADATEDVTAVERNPA 152

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:30:47 ; Search time 18 Seconds  
(without alignments)  
1261.255 Million cell updates/sec

Title: US-10-081-739a-2

Perfect score: 2459

Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQYGVMSYCGVG 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	484.5	19.7	494	1	AMY2_SALTY
2	472.5	19.2	514	1	AMY_EACAM
3	461.5	18.6	549	1	AMY_BACST
4	458.5	18.6	512	1	AMY_BACLI
5	454.5	18.5	518	1	AMY6_BACS7
6	447.5	18.2	495	1	AMY2_ECOLI
7	400	16.3	421	1	AMYA_VIGMU
8	396.5	16.1	443	1	AM2A_ORYSA
9	390.5	15.9	445	1	AMC2_ORYSA
10	366.5	14.9	427	1	AMY1_HORVU
11	365.5	14.8	438	1	AMY3_HORVU
12	363.5	14.8	368	1	AMY3_HORVU
13	357.5	14.5	437	1	AM3B_ORYSA
14	353	14.4	437	1	AM3C_ORYSA
15	351.5	14.3	429	1	AMY6_HORVU
16	347.5	14.1	440	1	AM3A_ORYSA
17	347	14.1	428	1	AM11_ORYSA
18	341.5	13.9	435	1	AM3D_ORYSA
19	336	13.7	437	1	AM3E_ORYSA
20	335.5	13.6	413	1	AMY3_WHEAT
21	284	11.5	548	1	AMT4_PSEST
22	278	11.3	551	1	AMT4_PSESA
23	263.5	10.3	710	1	CDGT_THETU
24	242	9.8	719	1	AMY1_BACST
25	236	9.6	464	1	AMY1_AERYH
26	235.5	9.6	711	1	CDGT_BACST
27	232.5	9.5	712	1	CDGT_BACS3
28	230.5	9.4	718	1	CDGT_BACS3
29	230	9.4	713	1	CDGT_BACSP
30	229	9.3	713	1	CDG2_PAENA
31	229	9.3	714	1	CDG1_PAENA
32	227	9.2	704	1	CDGT_BACOH
33	226	9.2	713	1	CDGT_BACSO

34	224.5	9.1	718	1	CDGT_BACLI
35	223.5	9.1	718	1	CDGT_BACCI
36	214.5	8.7	713	1	CDGU_BACCI
37	211	8.6	919	1	AMY_STRLI
38	210	8.5	703	1	CDGT_BACS2
39	205	8.3	1196	1	AMYB_PASPO
40	201.5	8.2	713	1	CDGT_BACS8
41	201	8.2	1481	1	APU_THEET
42	200	8.1	581	1	AMYI_SCHPO
43	198.5	8.1	498	1	AMYB_ASAPW
44	198.5	8.1	499	1	AMYB_ASAPW
45	196.5	8.0	499	1	AMY_ASPSH

ALIGNMENTS

RESULT 1  
AMY2\_SALTY  
ID AMY2\_SALTY STANDARD; PRT; 494 AA.  
AC P26613;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cytoplasmic alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).  
GN AMYA OR STM1963;  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
[1]  
RN RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=SW1103;  
RX MEDLINE=93015717; PubMed=1400215;  
RA Raha M., Kawagishi I., Mueller V., Kihara M., Macnab R.M.;  
RT "Escherichia coli produces a cytoplasmic alpha-amylase, AmyA.";  
RL J. Bacteriol. 174:6644-6652(1992).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=J72 / SGSC1412 / ATCC 700720;  
MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium L72.";  
RL Nature 413:852-856 (2001).  
[3]  
RP SEQUENCE OF 1-6 FROM N.A.  
RC STRAIN=J7M1103;  
MEDLINE=92407478; PubMed=1527488;  
RA Kawagishi I., Mueller V., Williams A.W., Irikura V.M., Macnab R.M.;  
RT "Subdivision of flagellar region III of the Escherichia coli and Salmonella typhimurium chromosomes and identification of two additional flagellar genes.";  
RL J. Gen. Microbiol. 138:1051-1065(1992).  
[4]  
RP SEQUENCE OF 476-494 FROM N.A.  
MEDLINE=93381452; PubMed=8371104;  
RA Raha M., Kihara M., Kawagishi I., Macnab R.M.;  
RT "Organization of the Escherichia coli and Salmonella typhimurium chromosomes between flagellar regions IIIa and IIb, including a large non-coding region.";  
RL J. Gen. Microbiol. 139:1401-1407(1993).  
CC -! CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.  
CC -! COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
CC -! SUBUNIT: Monomer (By similarity).  
CC -! SUBCELLULAR LOCATION: Cytoplasmic.  
CC -! SIMILARITY: Belongs to family 13 of glycosyl hydrolases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L01643; AAA27110.1; -;  
 CC EMBL; AS008787; AAL20875.1; -;  
 CC EMBL; M85241; AAA27079.1; -;  
 CC EMBL; L13280; AAA1970.1; -;  
 CC PIR; B45738; B45738.  
 CC HSP; P06278; 1VJS.  
 CC StyGene; SG10011; amva.  
 CC InterPro; IPR006589; Alp\_aml\_cat\_sub.  
 CC InterPro; IPR006047; Alpha\_aml\_cat.  
 CC Pfam; PF00128; alpha-amylase; 1.  
 CC SMART; SM00642; Amy; 1.  
 CC Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 CC Complete proteome.  
 CC ACT\_SITE 235 235 BY SIMILARITY.  
 CC ACT\_SITE 265 265 BY SIMILARITY.  
 CC ACT\_SITE 332 332 BY SIMILARITY.  
 CC METAL 104 104 CALCIUM (BY SIMILARITY).  
 CC METAL 239 239 CALCIUM (VIA CARBONYL OXYGEN) (BY  
 CC SIMILARITY).  
 CC CONFLICT 462 462 L -> S (IN REF. 1).  
 CC SEQUENCE 494 AA; 56522 MW; 5C1E863FED5E47C CRC64;  
 CC  
 CC Query Match 19.7%; Score 484.5; DB 1; Length 494;  
 CC Best Local Similarity 28.2%; Pred. No. 1.7e-27;  
 CC Matches 140; Conservative 74; Mismatches 160; Indels 123; Gaps 21;  
 CC  
 CC QY 9 KGVVQAFYWDVPSGIIWDTIRQIPEDYAGISAIIWIPASKMGGMGAYSGVDPYDF 68  
 CC DB 2 KNPTLLQYFWHYPDGGLKSELAEADGLDGINNWLPPACKGASGYSVGYDTYDL 61  
 CC QY 69 FDGEYDQKTVETFGSKQELNMINTAHAYCMKIADIVINHRAGGLEWNPFFV--- 124  
 CC DB 62 FDLGEFDQKGTIATKYGDKRLTAIDALKXNNIALLVNVNHNKMGDEKERIVQRVN 121  
 CC QY 125 -NDYT-----WTDFSKVA-SGKYTANYLDFH-----PNEHAGDSGTF- 160  
 CC DB 122 QDDTQTDNNIECEGWTRTFPARAGQYNSFTWYHCFSGIDHIENPDE----DGIFK 176  
 CC QY 161 -----GGYPDICHDK--SWD-----QYWL-WASQESYAAYLSIG 192  
 CC DB 177 IVNDYTGQWNDQVDEMGNFYILGNGENIDFRNHAVENTEIKYARWVMEQTHC- 229  
 CC QY 193 IDAWFDDYVGYAPVWVKDMLNMGWA-----VGEYDNTVDVAVLNWAYSSGAK--VF 244  
 CC DB 230 -DGRFLDAVXHPANFYKEMIEHVOAVAPKPLFVAEYWSHEVDKLOTYIDQVDGKTMLF 288  
 CC QY 245 DFALYYRMDEA-----FDKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTII- 294  
 CC DB 289 DAPLQMKFHEASRQAEYDMRHFT-----GTLVEADPFHATILVANHTDTPQALE 340  
 CC QY 295 -----WNKYPAAYAILIYE-QQTFIYRD-----YEEMLNK-----DKLNLLW 332  
 CC DB 341 APVEFPFKPLAYAILIRENGVPSVFPDLYGASVDESGNGETCTCRVDMVPVNLQRLIL 400  
 CC QY 333 IHENLAGGSTDIVYDNDDEL- FVRNGYGDKPGLIITYINLGSSKAGRWVYVPFAGACIH 391  
 CC DB 401 ARQFANG-IGTLFEDHPNCIAFSGSETEENFGCVVLSNGDGEKTLILGDNVANKTW 459  
 CC QY 392 EYTNLGGVNDKYVSS 408  
 CC DB 460 DFLGNR-----DEYVVTN 472

AMY\_BACAM  
 ID -AMY\_BACAM STANDARD; PRT; 514 AA.  
 AC P00692;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE glucanohydrolase).  
 OS Bacillus amyloliquefaciens  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OC NCBI\_TaxID=1390;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IH;  
 RX MEDLINE=83108808; PubMed=6185474;  
 RA Takkinen K., Pettersson R.F., Kalkkinen N., Palva I., Soederlund H.,  
 RA Kaeeriaeinen L.;  
 RA "Amino acid sequence of alpha-amylase from Bacillus amyloliquefaciens  
 RA deduced from the nucleotide sequence of the cloned gene.";  
 RL J. Biol. Chem. 258:1007-1013(1983).  
 RN (2)  
 RP SEQUENCE OF 32-222.  
 RX MEDLINE=80241725; PubMed=6156671;  
 RA Chung H.S., Friedberg F.;  
 RA "Sequence of the N-terminal half of Bacillus amyloliquefaciens alpha-  
 RA amylase.";  
 RL Biochem. J. 185:387-395(1980).  
 RN (3)  
 RP SEQUENCE OF 1-96 FROM N.A.  
 RX MEDLINE=82051296; PubMed=6170539;  
 RA Palva I., Pettersson R.F., Kalkkinen N., Lehtovaara P., Sarvas M.,  
 RA Soederlund H., Takkinen K., Kaeeriaeinen L.;  
 RA "Nucleotide sequence of the promoter and NH2-terminal signal peptide  
 RA region of the alpha-amylase gene from Bacillus amyloliquefaciens.";  
 RL Gene 15:43-51(1981).  
 RN (4)  
 RP SEQUENCE OF 1-39 FROM N.A.  
 RX MEDLINE=88137952; PubMed=2830166;  
 RA Ruohonen L., Hackman P., Lehtovaara P., Knowles J.K.C., Karaenen S.;  
 RA "Efficient secretion of Bacillus amyloliquefaciens alpha-amylase by  
 RA its own signal peptide from Saccharomyces cerevisiae host cells.";  
 RL Gene 59:161-170(1987).  
 RN (5)  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 32-331.  
 RX MEDLINE=20384196; PubMed=10924103;  
 RA Brzozowski A.M., Lawson D.M., Turkenburg J.P., Bisgaard-Prantzen H.,  
 RA Svendsen A., Borchert T.V., Dauter Z., Wilson K.S., Davies G.J.;  
 RA "Structural analysis of a chimeric bacterial alpha-amylase.  
 RA High-resolution analysis of native and ligand complexes.";  
 RL Biochemistry 39:9093-9107(2000).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -!- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J01542; AAA22191.1; -;  
 CC EMBL; V00092; CAA23430.1; -;  
 CC EMBL; A20154; CAA01489.1; -;  
 CC EMBL; M18424; AAA22192.1; -;  
 CC PIR; A92389; ALBSN.  
 CC PDB; 1E3X; 21-JUN-01.  
 CC PDB; 1E3Z; 24-JUN-03.  
 CC PDB; 1E40; 24-JUN-03.

DR PDB: 1E43; 21-JUN-01.  
DR InterPro: IPR006589; Alp\_amyl\_cat\_sub.  
DR InterPro: IPR006047; Alpha\_amyl\_cat.  
DR InterPro: IPR006046; Glyco\_hydro\_13.  
DR Pfam: PF00128; alpha-amylase; 1.  
DR PRINTS: PR00110; ALPHAAMYLASE.  
DR SMART: SM00642; Aamy; 1.  
KW Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
KW Signal; 3D-structure 31  
FT SIGNAL 1  
FT CHAIN 32 514  
FT ACT\_SITE 262 262  
FT ACT\_SITE 292 292  
FT ACT\_SITE 359 359  
FT METAL 133 133  
FT METAL 190 190  
FT METAL 212 212  
FT METAL 214 214  
FT METAL 225 225  
FT METAL 231 231  
FT METAL 233 233  
FT METAL 235 235  
FT METAL 266 266  
FT METAL 331 331  
FT METAL 438 438  
FT METAL 461 461  
FT CONFLICT 54 54  
FT CONFLICT 64 64  
FT CONFLICT 79 79  
FT CONFLICT 84 84  
SQ SEQUENCE 514 AA; 58403 MW; 3D566B3FB5CCDE7E CRC64;  
Query Match 19.2%; Score 472.5; DB 1; Length 514;  
Best Local Similarity 29.9%; Pred. No. 1.3e-26;  
Matches 147; Conservative 60; Mismatches 167; Indels 117; Gaps 22;  
QY 11 GVIMQAFVWDVSPGGIWDITRQKIPENYDAGISAIWIPPPASKMGGAYSGYDPYDFD 70  
DB 34 GTLMQYFEMYPNDQHQWKRQLDQAEHLSDIGITAVPIPPAYKGLSQS-DNGYGFYDIYD 92  
QY 71 LGEYDQKQTEVERFGSKOELVNMINTAHAYKVIAD-VINHRAGGD-----LEWNP- 122  
DB 93 LGEFOQKQITVTKYTKSELQDAIGSLSRNVQVGVGVVNLHKGADATEDVTAVEVNP 152  
QY 123 -----FVNDY---TWTFDSKVASGKYTANYLDHFHNLH----- 153  
DB 153 NRNOETSEYQIKAWTDFRFGNGN---TYSDFKWHVHPDGDWDSRKRISRFKPERGE 209  
QY 154 -----AGSGTGF--GYDICH-----KSHDQYWLWASQESYAAIYLSIGIDA 195  
DB 210 GKANDWEVSSNGNYDLYMADVDYDHPDVVAETKMG-----YANELSLDG 258  
QY 196 WRFDYVKGYPWVKDML-----NWMGMAVGEVWDTNVDVNLW--AYSSGAKVDFEA 247  
DB 259 FRIDAAKHIKFSFLRDWVQAVRQATKEMFTVAIYQWQNNAGKLENYLNKTSFNQSVFDP 318  
QY 248 LYYKDEA-----FDKNIPALVSLONGQTVVSRDPFKAVTFVANEHDT----- 292  
DB 319 LHFNLQASSGGGYDMR-----LLDG-TVVSRHPEKAVTFVENEHDTQPGQSLSTV 370  
QY 293 IIMNKYPAYAFILTYE-GOPTIFYRDY-----EELWKKKLNLIWIHENLAGGST 342  
DB 371 QTFKFLAYAFILTYREGYPQVFDGMVYGTGTFKPIPSLKONIEPILKARKEYAYGPQ 430  
QY 343 DIVYDNDLEI-FVRNGYD--KPLGITYINLGSSKAGRWVYVYVYVYVYVYVYVYVYVYV 399  
DB 431 H-DVIDHPDVIGTWREGDSSAAKGLAALITDGGGSKRMVYVYVYVYVYVYVYVYVYVYV 489  
QY 400 WVDKVVSSGW 410  
DB 490 TVK--IGSDGW 498

RESULT 3  
AMY\_BACST  
ID - AMY\_BACST STANDARD; PRT; 549 AA.  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
glucanohydrolase).  
GN AMYS.  
OS Bacillus stearothermophilus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.  
OX NCBI\_TaxID=1422;  
RW [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-39.  
RX MEDLINE=85234394; PubMed=3924897;  
RA Nakajima R., Imanaka T., Aiba S.;  
RT "Nucleotide sequence of the Bacillus stearothermophilus alpha-amylase  
gene.";  
RL J. Bacteriol. 163:401-406 (1985).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX STRAIN=DV5/PHI300;  
RA Ihara H., Sasaki T., Tsuboi A., Yamagata H., Tsukagoshi N., Uda S.;  
RT "Complete nucleotide sequence of a thermophilic alpha-amylase gene:  
homology between prokaryotic and eukaryotic alpha-amylases at the  
active sites.";  
RL J. Biochem. 98:95-103 (1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=NZ-3;  
RA MEDLINE=86159857; PubMed=3009417;  
RA Gray G.L., Mainzer S.E., Rey M.W., Lamsa M.H., Kindle K.L.,  
Carmona C., Requa C.;  
RT "Structural genes encoding the thermophilic alpha-amylases of  
Bacillus stearothermophilus and Bacillus licheniformis.";  
RL J. Bacteriol. 166:635-643 (1986).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Suominen I., Karp M., Lautamo J., Knowles J., Mantsaelae P.;  
RT "Thermophilic alpha amylase of Bacillus stearothermophilus: cloning,  
expression, and secretion by Escherichia coli.";  
RL (in) Chaloupka J., Krumphanz V. (eds.);  
RT Extracellular enzymes of microorganisms, pp.129-137, Plenum Press,  
New York (1987).  
RN [5]  
RP SEQUENCE OF 1-122 FROM N.A., AND SEQUENCE OF 35-48.  
RX STRAIN=DY-5;  
RA MEDLINE=86059211; PubMed=2999073;  
RA Tsukagoshi N., Iritani S., Sasaki T., Takemura T., Ihara H.,  
Igata Y., Yamagata H., Uda S.;  
RT "Efficient synthesis and secretion of a thermophilic alpha-amylase by  
protein-producing Bacillus brevis 47 carrying the Bacillus  
stearothermophilus amylase gene.";  
RL J. Bacteriol. 164:1182-1187 (1985).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=21125602; PubMed=11226887;  
RA Svay D., Fujimoto Z., Takase K., Matsumura M., Mizuno H.;  
RT "Crystal structure of Bacillus stearothermophilus alpha-amylase:  
possible factors determining the thermostability.";  
RL J. Biochem. 129:461-468 (2001).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
linkages in oligosaccharides and polysaccharides.  
CC -1- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.  
CC -1- SUBUNIT: Monomer.  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M11450; ARA22235.2; -  
 DR EMBL; X02769; CRA26547.1; -  
 DR EMBL; M57457; ARA22227.1; -  
 DR EMBL; M13255; ARA22241.1; -  
 DR PIR; A24436; A24436.  
 DR PIR; A91999; ALBSF.  
 DR PDB; 1HVX; 05-AUG-03.  
 DR InterPro; IPR006589; Alp\_aml cat sub.  
 DR InterPro; IPR006047; Alpha\_aml cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amyrase; 1.  
 DR PRINTS; PR00110; ALPHA-AMYLASE.  
 DR SMART; SM00642; Amy; 1.  
 DR Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 KW Signal; 3D-structure.  
 FT SIGNAL 1 34  
 FT CHAIN 35 549  
 FT ACT\_SITE 268 268  
 FT ACT\_SITE 272 272  
 FT ACT\_SITE 365 365  
 FT METAL 139 139  
 FT METAL 196 196  
 FT METAL 218 218  
 FT METAL 220 220  
 FT METAL 231 231  
 FT METAL 237 237  
 FT METAL 238 238  
 FT METAL 239 239  
 FT METAL 272 272  
 FT METAL 337 337  
 FT METAL 339 339  
 FT METAL 440 440  
 FT METAL 441 441  
 FT METAL 464 464  
 FT METAL 13 13  
 FT METAL 19 19  
 FT METAL 23 23  
 FT METAL 31 31  
 FT METAL 107 107  
 FT METAL 167 167  
 FT METAL 179 179  
 FT METAL 251 251  
 FT METAL 260 262  
 FT METAL 284 284  
 FT METAL 312 312  
 FT METAL 338 338  
 FT METAL 342 342  
 FT METAL 346 346  
 FT METAL 376 376  
 FT METAL 526 527  
 FT METAL 527 527  
 FT METAL 535 535  
 FT METAL 549 549  
 FT METAL 18.8; Score 461.5; DB 1; Length 549;  
 FT METAL 28.4; Pred. No. 8.6e-26;  
 FT METAL 146; Conservative 65; Mismatches 170; Indels 133; Gaps 25;  
 FT METAL 11 GVIMQAFVWVPSGGIWDITIRKIPEDWDAGISAINIPPASKMGAGSMGVDPVDFPD 70  
 FT METAL 40 GTTMCYFEWLPDDGTLTKVANEANNLSLIGITLWLPAYKGTGRS-DVGVGYVDLYD 98  
 FT METAL 71 LGEVDKQTVETRGSKOELVNMINTAHAYGMKVIADIVINHAGGD-LEW-----NPF 123  
 FT METAL 99 LGFENQKGAETKTKYTAQVLAQIAAAGAGMGMVADVDFDHKGADGTWDAVEVNP 158  
 FT METAL 124 VND-----YTWDFSKVASKGYTANYLDLPHNELHAG----- 155  
 FT METAL 159 DRNQEISGTQIAQWTKFDPPGRGN---TYSSPKRWYHFDGVDWDBSKRLSRVYKFRGI 215  
 FT METAL 215

Query Match  
 Best Local Similarity 18.8; Score 461.5; DB 1; Length 549;  
 Matches 146; Conservative 65; Mismatches 170; Indels 133; Gaps 25;  
 QY 11 GVIMQAFVWVPSGGIWDITIRKIPEDWDAGISAINIPPASKMGAGSMGVDPVDFPD 70  
 DB 40 GTTMCYFEWLPDDGTLTKVANEANNLSLIGITLWLPAYKGTGRS-DVGVGYVDLYD 98  
 QY 71 LGEVDKQTVETRGSKOELVNMINTAHAYGMKVIADIVINHAGGD-LEW-----NPF 123  
 DB 99 LGFENQKGAETKTKYTAQVLAQIAAAGAGMGMVADVDFDHKGADGTWDAVEVNP 158  
 QY 124 VND-----YTWDFSKVASKGYTANYLDLPHNELHAG----- 155  
 PB 159 DRNQEISGTQIAQWTKFDPPGRGN---TYSSPKRWYHFDGVDWDBSKRLSRVYKFRGI 215

QY 156 -----DSGTFG-----GYPDICHD-KSWDQYWLWASQESYAAYLRIGIDA 195  
 DB 216 GKAWDEVDTEGNYDYLMYADLMDHPVTVTELKSWGK-----YVNTNIDG 264  
 QY 196 WRDYKGYAPWVVKOLNWWG-----WAYGEYDNTVDVAVLWAYSSGA--KVDFPA 247  
 DB 265 FRLDAVAKHIFSPFDWLSVRSQTKPLFTVEGYWSYDINKHNYIMKNTGMTSLFDAP 324  
 QY 248 LYYKMDFA-----FDKNIPALVSALQNGQTVVSRDPFKAVTFVANHDT----- 292  
 DB 325 LHNKFTASKSGGTDFWRTLMT-----NTLMKDOPTLAVTFVDNHDTEPGQALQSWV 376  
 QY 293 IINKYKPAYAFILT-YEGOPTIEYRDE--EWLN-----XDKLKLNIWIHENLAGSTDIV 345  
 DB 377 DPWFKPLAYAFILTROEGYPCVFYGIPIQVNFSLASKIDPLLIARDVAYG-TQHD 435  
 QY 346 YYDNDELI-FVRNGYDGRP--GLITYINILGSSKAGRWVYVPK-FAGACHEYTNLGGWV 401  
 DB 436 YLDHSDIIGTWREGVTEKPGSLAALITDGP-GSKWTVGKHAGKVFYDLTGNSRDTV 494  
 QY 402 DKYVYSSGVWYLEAPAVDPANGY-----GYSVW 430  
 DB 495 T--INSDG-----GEFKVNGGSVSW 514  
 RESULT 4  
 ID AMY\_BACLI STANDARD; PRT; 512 AA.  
 AC P02778; Q84171;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Alpha-amyrase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE Glucanohydrolase) (BLA).  
 GN AMYS OR AMYL.  
 OS Bacillus licheniformis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1402;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27811;  
 RX MEDLINE=86111694; PubMed=2418011;  
 RA Yuuki T., Nomura T., Tezuka H., Tsuboi A., Yamagata H.,  
 RA Tsukagoshi N., Ueda S.;  
 RT "Complete nucleotide sequence of a gene coding for heat- and  
 RT pH-stable alpha-amyrase of Bacillus licheniformis: comparison of the  
 RT amino acid sequences of three bacterial liquefying alpha-amyases  
 RT deduced from the DNA sequences."  
 RL J. Biochem. 98:1147-1156(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86195857; PubMed=3009417;  
 RA Gray G.L., Mainzer S.E., Rey M.W., Lamsa M.H., Kindle K.L.,  
 RA Carmona C., Requaad C.;  
 RT "Structural genes encoding the thermophilic alpha-amyases of  
 RT Bacillus stearotherophilus and Bacillus licheniformis."  
 RL J. Bacteriol. 166:635-643(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84185455; PubMed=6609154;  
 RA Stephens M.A., Ortlepp S.A., Ollington J.F., McConnell D.J.;  
 RT "Nucleotide sequence of the 5' region of the Bacillus licheniformis  
 RT alpha-amyase gene: comparison with the B. amyloliquefaciens gene."  
 RL J. Bacteriol. 158:369-372(1984).  
 RN [5]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=89213924; PubMed=2540150;  
 RA Shalhoub M., Ziaei A.A., Ghaemi N., Pourbabaie A.A.;  
 RT "An unusual DNA sequence encoded a hyperthermostable alpha-amyase."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-104 FROM N.A.  
 RX MEDLINE=84185455; PubMed=6609154;  
 RA Stephens M.A., Ortlepp S.A., Ollington J.F., McConnell D.J.;  
 RT "Nucleotide sequence of the 5' region of the Bacillus licheniformis  
 RT alpha-amyase gene: comparison with the B. amyloliquefaciens gene."  
 RL J. Bacteriol. 158:369-372(1984).  
 RN [5]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE=89213924; PubMed=2540150;

RA Laioide B.M., Chambliss G.H., McConnell D.J.;  
 RT "Bacillus licheniformis alpha-amylase gene, amyL, is subject to  
 RT promoter-independent catabolite repression in *Bacillus subtilis*.";   
 RL J. Bacteriol. 171:2435-2442(1989).  
 RN [6]  
 RP SEQUENCE OF 30-47.  
 RX MEDLINE=82098050; PubMed=6172418;  
 RA Kuhn H., Fietzek P.P., Lampen J.O.;  
 RT "N-terminal amino acid sequence of *Bacillus licheniformis*  
 RT alpha-amylase: comparison with *Bacillus amyloliquefaciens* and  
 RL *Bacillus subtilis* enzymes.";   
 RL J. Bacteriol. 149:372-373(1982).  
 RN [7]  
 RP MAPPING OF SUBSTRATE-BINDING SITE.  
 RX MEDLINE=2192788; PubMed=11997021;  
 RA Kandra L., Gyemant G., Remenyik J., Hovanszki G., Liptak A.;  
 RT "Action pattern and sub-site mapping of *Bacillus licheniformis*  
 RT alpha-amylase (BLA) with modified maltooligosaccharide substrates.";   
 RL FEBS Lett. 518:79-82(2002).  
 RN [8]  
 RP MUTAGENESIS OF HIS-64; HIS-162; HIS-276; HIS-322; HIS-435 AND HIS-479.  
 RC STRAIN=ATCC 6598;  
 RX MEDLINE=90368748; PubMed=2394736;  
 RA Declerck N., Joyet P., Gaillardin C., Masson J.M.;  
 RT "Use of amber suppressors to investigate the thermostability of  
 RT *Bacillus licheniformis* alpha-amylase. Amino acid replacements at 6  
 RT histidine residues reveal a critical position at His-133.";   
 RL J. Biol. Chem. 265:15481-15488(1990).  
 RN [9]  
 RP MUTAGENESIS OF ALA-238.  
 RC STRAIN=ATCC 6598;  
 RX MEDLINE=96367070; PubMed=8771184;  
 RA Declerck N., Joyet P., Trosset J.Y., Garnier J., Gaillardin C.;  
 RT "Hyperthermostable mutants of *Bacillus licheniformis* alpha-amylase:  
 RT multiple amino acid replacements and molecular modelling.";   
 RL Protein Eng. 8:1029-1037(1995).  
 RN [10]  
 RP MUTAGENESIS OF ASP-150; ASN-155; ARG-175; ASP-193; ASN-201; GLN-207;  
 RP ASN-217; ASN-219; ASN-221; ASP-229; ASP-233; ALA-298; GLU-300; GLN-359  
 RP AND GLU-365.  
 RC STRAIN=ATCC 6598;  
 RX MEDLINE=20425100; PubMed=10966804;  
 RA Declerck N., Machius M., Wiegand G., Huber R., Gaillardin C.;  
 RT "Probing structural determinants specifying high thermostability in  
 RT *Bacillus licheniformis* alpha-amylase.";   
 RL J. Mol. Biol. 301:1041-1057(2000).  
 RN [11]  
 RP MUTAGENESIS OF GLN-293 AND ASN-294.  
 RC STRAIN=ATCC 6598;  
 RX MEDLINE=22622182; PubMed=12736372;  
 RA Declerck N., Machius M., Joyet P., Wiegand G., Huber R.,  
 RA Gaillardin C.;  
 RT "Hyperthermostabilization of *Bacillus licheniformis* alpha-amylase and  
 RT modulation of its stability over a 50 degrees C temperature range.";   
 RL Protein Eng. 16:287-293(2003).  
 RN [12]  
 RP MUTAGENESIS OF TRP-292 AND VAL-315.  
 RC STRAIN=ATCC 27811;  
 RX MEDLINE=22797417; PubMed=12915728;  
 RA Rivera M.H., Lopez-Munguia A., Soberon X., Saab-Rincon G.;  
 RT "Alpha-amylase from *Bacillus licheniformis* mutants near to the  
 RT catalytic site: effects on hydrolytic and transglycosylation  
 RT activity.";   
 RL Protein Eng. 16:505-514(2003).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RC STRAIN=ATCC 27811;  
 RX MEDLINE=95182462; PubMed=7877175;  
 RA Machius M., Wiegand G., Huber R.;  
 RT "Crystal structure of calcium-depleted *Bacillus licheniformis* alpha-  
 RT amylase at 2.2-A resolution.";   
 RL J. Mol. Biol. 246:545-559(1995).  
 RN [14]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=98212915; PubMed=9551551;  
 RA Machius M., Declerck N., Huber R., Wiegand G.;  
 RT "Activation of *Bacillus licheniformis* alpha-amylase through a  
 RT disorder-to-order transition of the substrate-binding site mediated  
 RT by a calcium-sodium-calcium metal triad.";   
 RL Structure 6:281-292(1998).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 330-512.  
 RX MEDLINE=20384196; PubMed=10924103;  
 RA Brzozowski A.M., Lawson D.M., Turkenburg J.P., Bisgaard-Prantzen H.,  
 RA Svendsen A., Borchert T.V., Dauter Z., Wilson K.S., Davies G.J.;  
 RT "Structural analysis of a chimeric bacterial alpha-amylase.  
 RT High-resolution analysis of native and ligand complexes.";   
 RL Biochemistry 39:9099-9107(2000).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF MUTANT  
 RP H162V/N219F/A238V/Q293S/N294Y.  
 RC STRAIN=ATCC 6598;  
 RX MEDLINE=22538505; PubMed=12540849;  
 RA Machius M., Declerck N., Huber R., Wiegand G.;  
 RT "Kinetic stabilization of *Bacillus licheniformis* alpha-amylase through  
 RT introduction of hydrophobic residues at the surface.";   
 RL J. Biol. Chem. 278:11546-11553(2003).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- BIOTECHNOLOGY: Used in the food industry for high temperature  
 CC liquefaction of starch-containing mash and in the detergent  
 CC industry to remove starch. Sold under the name Termamyl by  
 CC Novozymes.  
 CC -1- MISCELLANEOUS: Able to work at relatively high (alkaline) pH  
 CC values (up to pH 11) and at high temperatures (up to 100 degrees  
 CC Celsius).  
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; X03236; CAA26981.1; -  
 CC EMBL; M38570; AAA22226.1; -  
 CC EMBL; M3256; AAA22240.1; -  
 CC EMBL; K01984; AAA22193.1; -  
 CC EMBL; AF438149; AAO26743.1; -  
 CC EMBL; M26412; AAA22237.1; -  
 CC EMBL; AL7930; CAA01355.1; -  
 CC PIR; A91997; AUBSU.  
 CC DR PDB; 1BLI; 23-MAR-99.  
 CC DR PDB; 1BPL; 17-AUG-96.  
 CC DR PDB; 1E3X; 21-JUN-01.  
 CC DR PDB; 1E3Z; 24-JUN-03.  
 CC DR PDB; 1E40; 24-JUN-03.  
 CC DR PDB; 1E43; 21-JUN-01.  
 CC DR PDB; 1O80; 03-APR-03.  
 CC DR PDB; 1VUS; 12-MAR-97.  
 CC DR InterPro; IPR006589; Alp\_aml\_cat\_sub.  
 CC DR InterPro; IPR006047; Alpha\_aml\_cat.  
 CC DR InterPro; IPR006046; Glyco\_hydro\_13.  
 CC DR Pfam; PF00128; alpha-amylase; 1.  
 CC DR PRINTS; PR00110; ALPHAMYLASE.  
 CC DR SMART; SM00642; Amy; 1.  
 CC DR Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 CC Signal; 3D-structure.  
 CC FT SIGNAL 1 29  
 CC FT CHAIN 30 512 ALPHA-AMYLASE.  
 CC FT ACT\_SITE 260 260  
 CC FT ACT\_SITE 264 264



```

FT ACT SITE 357 357
FT METAL 133 133
FT METAL 190 190
FT METAL 210 210
FT METAL 212 212
FT METAL 223 223
FT METAL 229 229
FT METAL 231 231
FT METAL 233 233

Query Match 18.6%; Score 458.5; DB 1; Length 512;
Best Local Similarity 28.6%; Pred. No. 1.3e-25; Indels 109; Gaps 17;
Matches 139; Conservative 68; Mismatches 170;

QY 11 GVIMQAFYWDVPSGGIWMWDTIRQKIPWYDAGISAIWIPPPASKMGAYSMGYDYPDFD 70
Db 34 GTLMQYFEWYMPDQHWKRLQNDLSAYLAEGHGTAVWIPPAKGTQSA-DVGYGAYDLYD 92
QY 71 LGEYDQKGTETRFSGKQELVNMINTAHAYCMKVIADIVLNHRAAGDLEWNPVNDYWT 130
Db 93 LGEFHQKGTATKYTGKELQSAKLSHSDINVYGDVWLNHKGADATEDVTAVEVDP 152
QY 131 DFSKVASGKYT-ANYLDFHPNELH- 153
Db 153 DNRVVISGEHRIKAWTHFHPGRGSTYDFKWHYHFDGTDWDESKLNRIYKFGKAWD 212
QY 154 --AGDSCTFG--CYPDICH- --KSDQYWLWASQESYAAAYLSIGIDAWRPD 200
Db 213 WEVSNENGYDLYMADIDYDHPDVAEIKRWGTW- --YANELQDQGRFLDA 261
QY 201 VKGYAPWVYKDLNWWGG- --WAYGEYDNTNDAVLNWAYSS- --GAKVDFDLYYK 252
Db 262 VKHIFESLRDVAHAREKTKGEMFTVAEYKQNDLGALENLYKTNFNSVDFVPLHYQF 321
QY 253 DEA- --FDKNIIPALVSALQNGQVVSRRDPKATFVANHDT- --IINWK 297
Db 332 HAASQGGYDMRKL- --LNSTVSKHPLKAVTFVDNHDTPQGSLSSTVQWTFK 373
QY 298 YPAAVAFILTYE-GOFTIFYRDY- --EWMKNDKLNLIWHENLAGSDTDVY 347
Db 374 PLAYAFILTRSGYPOVYFGDMYTKGDSQREIPALKHKEIPILKARKQYAVGAQH-DYF 432
QY 348 DNEDEL-FVRNGYGD--KPLGITVINLSSKAGRWVYVPEACACHETVGNLGGWVDY 404
Db 433 DHHDIVGTREGDSSVANSGLAIDTGPQGAARMYVGRQAGETWHDITGNRSEPV-V 490
QY 405 VYSSGW 410
Db 491 INSEGW 496

RESULT 5
AMT6_BACS7
ID AMT6_BACS7 STANDARD; PRT; 518 AA.
AC P19571;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glucan 1,4-alpha-maltohexaosidase precursor (EC 3.2.1.98) (G6-amyase)
DE (Maltotetraose-producing amylase) (Exo-maltohexaolylase).
OS Bacillus sp. (strain 707).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1416;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 34-36.
RX MEDLINE=88162814; PubMed=3258152;
RA Tsukamoto A., Kimura K., Ishii Y., Takano T., Yamane K.;
RT "Nucleotide sequence of the maltotetraose-producing amylase gene from
RT an alkalophilic Bacillus sp. #707 and structural similarity to
RL liquefying type alpha-amyases."
RL Biochem. Biophys. Res. Commun. 151:25-31(1988).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-D-glucosidic linkages
CC in amylaceous polysaccharides so as to remove successive

```

```

CC maltotetraose residues from the non-reducing chain ends.
CC -!- COFACTOR: Binds 2 calcium ions and 1 sodium ion per subunit (By
CC similarity).
CC -!- PATHWAY: Starch degradation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M18862; AAA22231.1; -.
CC PIR; A27705; A27705.
CC HSP; P06278; 1VJ5.
CC InterPro; IPR006589; Alp_amy1_cat_sub.
CC InterPro; IPR006047; Alpha_amy1_cat.
CC InterPro; IPR006046; Glyco_Hydro_13.
CC Pfam; PF00128; alpha-amyase; 1.
CC PRINTS; PR00110; ALPHAAMYLASE.
CC SMART; SM00642; Amy; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Signal.
FT SIGNAL 34 33
FT CHAIN 1 518 GLUCAN 1,4-ALPHA-MALTOHEXAOSIDASE.
FT ACT_SITE 269 269 BY SIMILARITY.
FT ACT_SITE 273 273 BY SIMILARITY.
FT ACT_SITE 366 366 BY SIMILARITY.
FT METAL 139 139 CALCIUM 1 (BY SIMILARITY).
FT METAL 196 196 CALCIUM 2 AND SODIUM (BY SIMILARITY).
FT METAL 219 219 CALCIUM 2 (VIA CARBONYL OXYGEN) (BY
FT METAL 221 221 SIMILARITY).
FT METAL 232 232 CALCIUM 2 AND SODIUM (BY SIMILARITY).
FT METAL 238 238 CALCIUM 1 AND SODIUM (BY SIMILARITY).
FT METAL 240 240 CALCIUM 2 (BY SIMILARITY).
FT METAL 242 242 CALCIUM 2 (BY SIMILARITY).
FT METAL 273 273 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY
FT METAL 273 273 SIMILARITY).
SQ SEQUENCE 518 AA; 59009 MW; 3A961E21612682C4 CRC64;
Query Match 18.5%; Score 454.5; DB 1; Length 518;
Best Local Similarity 29.1%; Pred. No. 2.5e-25;
Matches 141; Conservative 71; Mismatches 166; Indels 107; Gaps 23;
QY 11 GVIMQAFYWDVPSGGIWMWDTIRQKIPWYDAGISAIWIPPPASKMGAYSMGYDYPDFD 70
Db 40 GTLMQYFEWYMPDQHWKRLQNDLSAYLAEGHGTAVWIPPAKGTQSA-DVGYGAYDLYD 98
QY 71 LGEYDQKGTETRFSGKQELVNMINTAHAYCMKVIADIVLNHRAAGD- --LEWNP- 122
Db 99 LGEFNQKGTATKYTGKELQSAKLSHSDINVYGDVWLNHKGADATEDVTAVEVDP 158
QY 123 - --FVNDYT- --WTFDFSKVASGK- --YTANYLDFHPNELHAGSDGTGGYD 165
Db 159 NRQEVGTGTYEATWTRFDPPGRGNTHSSFKWYHFDGVDWDQSRLLNNRIYKFG- -- 215
QY 166 ICHDKSWD- --QYWLWAS- --QESYAAAYLSIGIDAWRPDYK 203
Db 216 - --HGRKAWDEVDTENGYDLYMADIDTGPQGAARMYVGRQAGETWHDITGNRSEPV-V 273
QY 204 YAFWVWKDLNWWGG- --WAVGEYDNTNDAV- --LNWAYSSGAKVDFDLYYK 251
Db 274 IKYSFTEDLNHVSATGKNMFAVAFKNDJGAENVLQKNWHS- --VFDVPLHN 329
QY 252 MDEA- --FDKNIIPALVSALQNGQVVSRRDPKATFVANHDT- --IINW 296
Db 330 LYNASKSGGNYDMRNT- --FNG-TVQSHPHAVTFVDNHDSDQPEALESEVEWF 381
QY 297 KYPAYAFILTYE-GOFTIFYRDY- --LNKDKLNLIWHENLAGSDTDIVYDN 349

```





QY 337 LAGGSTDIVYNDDEL-FVRNGYCKPGLITYINLG 372  
DB 405 FAHG-VQTLFFDPHNCIAFSRSGTDFEFCGVVMSNG 440

RESULT 7  
AMYA\_VIGMU  
ID AMYA\_VIGMU STANDARD; PRT; 421 AA.  
AC P17859;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
glucanohydrolase).  
GN AMY1.1.  
OS Vigna mungo (Rice bean) (Black gram).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurasids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.  
OX NCBI\_TaxID=3915;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cotyledon;  
RX MEDLINE=90332425; PubMed=2377468;  
RA Yamauchi D., Minamikawa T.;  
RT "Nucleotide sequence of cDNA for alpha-amylase from cotyledons of  
germinating Vigna mungo seeds".  
RL Nucleic Acids Res. 18:4250-4250(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94120017; PubMed=8290640;  
RA Takeuchi H., Yamauchi D., Wada S., Minamikawa T.;  
RT "Nucleotide sequence of the alpha-amylase gene from Vigna mungo".;  
RL Plant Physiol. 103:1459-1459(1993).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
linkages in oligosaccharides and polysaccharides.  
CC -1- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53049; CAA37217.1; -;  
CC EMBL; X73301; CAA51734.1; -;  
CC PIR; S10514; S10514.  
CC HSP; P04063; IAVA.  
CC InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
CC InterPro; IPR008047; Alp\_amy1\_cat.  
CC InterPro; IPR008046; Glyco\_hydro\_13.  
CC Pfam; PF00128; alpha-amylase; 1.  
CC PRINTS; PR00110; ALPHAMYLASE.  
CC SMART; SM00642; Amy; 1.  
CC Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
KW Signal.  
FT SIGNAL 1 23 PROBABLE.  
FT CHAIN 24 421 ALPHA-AMYLASE.  
FT ACT\_SITE 201 201 BY SIMILARITY.  
FT ACT\_SITE 309 309 BY SIMILARITY.  
FT METAL 113 113 CALCIUM 1 (BY SIMILARITY).  
FT METAL 130 130 CALCIUM 2 (BY SIMILARITY).  
FT METAL 133 133 CALCIUM 2 (BY SIMILARITY).  
FT METAL 135 135 CALCIUM 2 (BY SIMILARITY).  
FT METAL 139 139 CALCIUM 3 (BY SIMILARITY).  
FT METAL 149 149 CALCIUM 3 (BY SIMILARITY).  
FT METAL 160 160 CALCIUM 3 (BY SIMILARITY).  
FT METAL 168 168 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY  
SIMILARITY).

FT METAL 170 170 CALCIUM 1 AND 3 (BY SIMILARITY).  
SQ SEQUENCE 421 AA; 46888 MW; 15CA0DABA3DB4656 CRC64;  
Query Match 16.3%; Score 400; DB 1; Length 421;  
Best Local Similarity 28.6%; Pred. No. 1.6e-21;  
Matches 118; Conservative 57; Mismatches 109; Indels 128; Gaps 19;  
QY 12 VIMQAFVWDVPSGGIWMDTIRQKIPWYDAGISAIWIPPPASKMGSGAYSMGYDYPDFDL 71  
DB 25 LIFQGFNWESSKGGWNSLKNSIPDLANAGITHVWLPSPQSVS---PEGYLPGRLYDL 81  
QY 72 GEYDQGTVETREFGSKOELVNMINTAHAYGMKVIADIVINHEA-----GGD 117  
DB 82 D-----ASKYGSKNELKSLAFAHFKIGIKCLADIVINHRTAERKDGKRGYICIFEGT 133  
QY 118 ----LEWNPVNDYTWTFDSKVASGKYTANYLDFHFNELHAGDSGTFFGY---PDICH-- 168  
DB 134 PDSRQMDGP---SFICRDDTAYSDG--TGN-----NDSGE--GYDAAPDIDHLN 175  
QY 169 ---DKSMDQYWLWASQESYAAVLRISIGIDARPDYKGYAPWVVKWLNW--GGWAVEY 224  
DB 176 PQVQRELSWMLKTE-----IGFDGWRDFVKGYAPSIKTYMEQTKPDFAVGSK 227  
QY 225 WDT-----NVD-----AVLNWYSSGAKVDFDALVYKMDAEPDNKNI PALVSALQ 269  
DB 228 WDSISYQDQKPNYNQSHRGALVNWVESAGGAI-----TAFDFTTKGILQANQ 277  
QY 270 -----NGQT--VVSRRDPFKAVTFVANHDTDI--WNKYP-----AYAFILTYEGQ 310  
DB 278 GELWRLIDPNKPKGMIGVKNPENAFTVDNHDGTSTQRLWPPSPDKVMQGYAVILTHPGT 337  
QY 311 PTIPYRDEEWLNKDKLNLIWHENLAGGSTDIVYDNDDELIFVRNGYDK 362  
DB 338 PSIFYDFHFDWGLKEQIAKL-----SSIRLRNGINEK 369  
RESULT 8  
ID AMYA\_ORYSA STANDARD; PRT; 443 AA.  
AC P27935;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Alpha-amylase isozyme 2A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
glucanohydrolase).  
GN AMY1.5 OR AMY2A.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=cv. Japonica M202;  
RC MEDLINE=92175526; PubMed=1541400;  
RA Huang N., Reini S.J., Rodriguez R.L.;  
RT "AMY2A: a novel alpha-amylase-encoding gene in rice.";  
RL Gene 111:223-228(1992).  
CC -1- FUNCTION: Important for breakdown of endosperm starch during  
germination.  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
linkages in oligosaccharides and polysaccharides.  
CC -1- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
CC -1- SUBUNIT: Monomer.  
CC -1- DEVELOPMENTAL STAGE: Expressed at a high level during germination  
in the aleurones cells under the control of the plant hormone  
gibberellic acid and in the developing grains at a low level.  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way



QY 12 VIMQAFYNDV--PSGGIWNVTIROKIPEDYWDAGISAIWIPASKMGCGAYSMGYDPYDF 69  
 Db 24 ILFOGFWESWRQSGG-WYNLLMGKVDIVAAAGVTHVWLPFPHSVS---TQYMPGRLY 79  
 QY 70 DLGEYDQKGTVEYTRFGSKOELVNNMINTAHAYGMKVIADIVINHRAG-----115  
 Db 80 DLD-----ASRYGTSMLKSLISALHGKGIQAIADVVINHRCADYKDSRGYCIPEG 131  
 QY 116 ---GDLFWNFVNDYTWDFSKVASKYTYANVLDHFNELHAGDSGT-FGGYDICH-- 168  
 Db 132 GTPDRLGDMGHMCRDTPQS-----DGTGNL-----DTGADFAAPDIDHUN 175  
 QY 169 ---DKSWDQYWLWASQESYAYLRSIGIDAWREDYVKGYPWVVDLWNNWG--GWAVE 223  
 Db 176 GVQRELTDWLLMLKSD-----EVGFDAWLDLPARGYSPEVAKVIEGTTFVGLAVAE 228  
 QY 224 YNDT-----NVD-----AVLW-----AYSSGAKVDFALYKMDFAEDKNVIPA 263  
 Db 229 LWDNSVAYGGDKPEYNQDAHQALVDWDRVGGTASAGWVDFDTTKGINVAVEGB-LWR 287  
 QY 264 LVSAALQNGQTVSRDPPKAVTFVANHTDIIWKNIP-----AYAFILTYEGOPTIFY 315  
 Db 288 LIDQGRAPGVIGWPAKAVTFVDNHTDGTQGWPPSPDKWQGVAYILTHPGNCFIFY 347  
 QY 316 RDEEWLNLKDLKXLIWHEN---LAGSTDIYVYNDDEL 353  
 Db 348 DHFFDWGLKQIQAALVAVRQRNGVTATSSUKIMLHDADAYV 388

## RESULT 10

AMV2 HORVU STANDARD; PRT; 427 AA.  
 AC P04063;  
 DT 01-NOV-1986 (Rel. 03, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-  
 DE glucan glucanohydrolase) (AMV2-2) (High pi alpha-amylase).  
 GN AMV1.2  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rahmatullah R.J., Huang J.K., Clark K.L., Reesk G.R.,  
 RA Chandra G.R., Muthukrishnan S.;  
 RA "Nucleotide and predicted amino acid sequences of two different genes  
 RT for high-pi alpha-amylases from barley.";  
 RL Plant Mol. Biol. 12:119-121(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85131184; PubMed=3871776;  
 RA Rogers J.C.;  
 RA "Two barley alpha-amylase gene families are regulated differently in  
 RT aleurone cells.";  
 RL J. Biol. Chem. 260:3731-3738(1985).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=94254083; PubMed=8196040;  
 RA Kadziola A., Abe J.-I., Svensson B., Haser R.;  
 RA "Crystal and molecular structure of barley alpha-amylase.";  
 RL J. Mol. Biol. 239:104-121(1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH BASI.  
 RC STRAIN=cv. Mennet;  
 RX MEDLINE=98298441; PubMed=9634702;  
 RA Vallee F., Kadziola A., Bourne Y., Juy M., Rodenburg K.W.,  
 RA Svensson B., Haser R.;  
 RA "Barley alpha-amylase bound to its endogenous protein inhibitor BASI:  
 RT crystal structure of the complex at 1.9-A resolution.";  
 RL Structure 6:649-659(1998).

CC -- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -- COFACTOR: Binds 3 calcium ions per subunit.  
 CC -- SUBUNIT: Monomer.  
 CC -- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally  
 CC regulated. Germinating embryos produce the hormone gibberellic  
 CC acid, which within 10 hours stimulates the aleurone cells covering  
 CC the endosperm of the seed to produce alpha-amylase. The enzyme  
 CC then degrades the starch within the endosperm for use by the  
 CC developing plant embryo.  
 CC -- INDUCTION: Type B isozyme mRNA is undetectable in unstimulated  
 CC cells and increases a hundred-fold after stimulation with  
 CC gibberellic acid.  
 CC -- MISCELLANEOUS: There are at least 4 types of alpha-amylase in  
 CC barley.  
 CC -- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X15226; CAA33298.1; -;  
 CC EMBL; K02637; AAA98790.1; -;  
 CC PIR; A31960; ALEHB.  
 CC PDB; 1AMY; 13-MAY-95.  
 CC PDB; 1AVA; 16-MAR-99.  
 CC PDB; 1BG9; 15-JUN-99.  
 CC InterPro; IPR006589; Alp\_amyl\_cat.sub.  
 CC InterPro; IPR006047; Alpha\_amyl\_cat.  
 CC InterPro; IPR006046; Glyco\_Hydro\_13.  
 CC Pfam; PF00128; alpha-amylase; 1.  
 CC PRINTS; PR00110; ALPHAAMYLASE.  
 CC SMART; SM00642; Amy; 1.  
 CC Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
 CC Calcium-binding; signal; Multigene family; 3D-structure.  
 CC SIGNAL 1 24 ALPHA-AMYLASE TYPE B ISOZYME.  
 FT CHAIN 25 427  
 FT ACT\_SITE 203 203  
 FT ACT\_SITE 228 228  
 FT ACT\_SITE 313 313  
 FT METAL 115 115  
 FT METAL 132 132  
 FT METAL 135 135  
 FT METAL 137 137  
 FT METAL 141 141  
 FT METAL 151 151  
 FT METAL 162 162  
 FT METAL 165 165  
 FT METAL 166 166  
 FT METAL 167 167  
 FT METAL 170 170  
 FT METAL 172 172  
 FT CONFLICT 134 134  
 FT CONFLICT 195 197  
 FT CONFLICT 425 425  
 FT STRAND 27 29  
 FT TURN 33 34  
 FT HELIX 35 37  
 FT TURN 39 40  
 FT HELIX 42 47  
 FT TURN 48 49  
 FT HELIX 50 56  
 FT TURN 57 57  
 FT STRAND 60 63  
 FT TURN 69 69  
 FT STRAND 72 73  
 FT TURN 77 77  
 FT TURN 80 81  
 FT TURN 84 85  
 G -> D (IN REF. 1).  
 IGF -> HRL (IN REF. 2).  
 E -> Q (IN REF. 1).

PT TURN 87 88  
 FT HELIX 91 104  
 FT TURN 105 105  
 FT STRAND 107 112  
 FT STRAND 116 116  
 FT STRAND 121 122  
 FT TURN 124 125  
 FT STRAND 128 130  
 FT TURN 139 140  
 FT HELIX 144 146  
 FT STRAND 147 147  
 FT TURN 149 150  
 FT TURN 152 154  
 FT STRAND 165 165  
 FT TURN 168 169  
 FT STRAND 172 173  
 FT TURN 175 176  
 FT HELIX 178 193  
 FT TURN 194 195  
 FT STRAND 199 202  
 FT TURN 203 204  
 FT HELIX 205 207  
 FT STRAND 210 220  
 FT STRAND 224 227  
 FT STRAND 235 235  
 FT TURN 237 238  
 FT STRAND 241 241  
 FT HELIX 246 260  
 FT TURN 261 262  
 FT STRAND 265 268  
 FT TURN 270 279  
 FT HELIX 280 282  
 FT TURN 284 287  
 FT TURN 290 291  
 FT HELIX 297 299  
 FT TURN 300 300  
 FT HELIX 302 304  
 FT STRAND 305 308  
 FT TURN 312 314  
 FT TURN 316 318  
 FT HELIX 325 327  
 FT HELIX 328 337  
 FT STRAND 341 345  
 FT HELIX 346 350  
 FT TURN 351 351  
 FT HELIX 355 367  
 FT TURN 368 369  
 FT TURN 372 373  
 FT STRAND 376 382  
 FT TURN 383 384  
 FT STRAND 385 390  
 FT TURN 391 393  
 FT STRAND 394 398  
 FT HELIX 405 407  
 FT STRAND 412 418  
 FT TURN 419 420  
 FT STRAND 421 427  
 SQ SEQUENCE 427 AA; 47355 MW; 957C0B16621BF748 CRC64;

Query Match 14.9%; Score 366.5; DB 1; Length 427;  
 Best Local Similarity 26.8%; Pred. No. 4.1e-19;  
 Matches 114; Conservative 60; Mismatches 132; Indels 123; Gaps 19;

QY 5 SELEKGGVIMQAFYVDV-PSGGIWDITROKIPWYDAGISAIWIPPASKGNGGAYSNGY 63  
 DB 19 ASLASQVLPGFNWESKXNGWYFLGKYDDIAAGITHVLPFPASQSV-EGY 75  
 QY 64 DPYDFDLGDKYDQGVTRFGSKQBLVNMINTAHYGNKVIADIVINHRA----- 114  
 DB 76 MPCRLYDL-----ASKYGNKAQLKSLGALHGKGVKATADIVINHRTAEHKDGRGI 127  
 QY 115 -----LEWNPV-----NDYTWTFPSKVASKGYTANYLDFHNPHELHAGDSGT-FG 161

DB 128 YCIFEGGTPDARLDWGNPHMICRDRPYAD-----GTGN-----PTGADFG 168  
 QY 162 GYPDICH-----DKSWDQVWMLWASQESYAAVLRISGIDAMRFDYVYKGYAPVWVKWMLN-W 215  
 DB 169 AAPDIDHLNLRVQKELVWLNWKAD-----IGPDGWRFDPAKGYSDADVAKIYIDRS 220  
 QY 216 WGVNAVGEYW-----DTNVDVNLWAYSSGAK-----VDFPALYKMDFAF 256  
 DB 221 EFSFAVAETWTSLAYGGDKPNLNQDHRQLVNMVVDKVGKGPATTFDFTTKGILNAV 280  
 QY 257 DNK-----NIPALYSALQNGQTVVSRDPPFKAFTFVANHDTDIWNKYP----- 299  
 DB 281 EGEINRLRGTDCKAKPGMIGW-----PAKATFVDNHDGSGTCHMPPFSDRVMQ 330  
 QY 300 AYAFILTYEGQTFIYRDYEWLNKDKLNLI-----WIHE-----NLGSGSTIVVYDN 349  
 DB 331 GYAYILTHPGTCIFDHFDDGLKKEIDRLVSVTRHGIHNEKSKLOITIEADADLYLAEI 390  
 QY 350 DELIFVRNG 358  
 DB 391 DGKVIKLG 399

RESULT 11  
 AMYL\_HORVU  
 ID AMYL\_HORVU STANDARD; PRT: 438 AA.  
 AC P00633;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type A isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase) (AMYL1) (Low pI alpha-amylase).  
 GN AMYL1.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum.  
 OC NCBI\_TaxID:4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Himalaya;  
 RX MEDLINE=83238423; PubMed=6190808;  
 RA Rogers J.C., Millman C.;  
 RT "Isolation and sequence analysis of a barley alpha-amylase cDNA clone.";  
 RL J. Biol. Chem. 258:8169-8174(1983).  
 CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.  
 CC -|- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC -|- SUBUNIT: Monomer (By similarity).  
 CC -|- SUBCELLULAR LOCATION: Extracellular.  
 CC -|- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally regulated. Germinating embryos produce the hormone gibberellic acid, which within 10 hours stimulates the aleurone cells covering the endosperm of the seed to produce alpha-amylase. The enzyme then degrades the starch within the endosperm for use by the developing plant embryo.  
 CC -|- MISCELLANEOUS: There are at least 4 types of alpha-amylase in barley.  
 CC -|- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J01236; AAA32929.1; -  
 CC PIR: A00846; ALEH.  
 CC HSSP: P04063; IAVA.  
 CC InterPro: IPR006589; Alp\_amyl\_cat\_sub.

DR InterPro: IPR006047; Alpha amyl cat.  
 DR InterPro: IPR006046; Glyco\_hydro\_13.  
 DR Pfam: PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SMO0642; Amy; 1.  
 DR Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
 KW Calcium-binding; Signal; Multigene family.  
 FT SIGNAL 1 24  
 FT CHAIN 25 438 ALPHA-AMYLASE TYPE A ISOZYME.  
 FT ACT\_SITE 204 204 BY SIMILARITY.  
 FT ACT\_SITE 229 229 BY SIMILARITY.  
 FT ACT\_SITE 315 315 BY SIMILARITY.  
 FT METAL 116 116 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 133 133 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 136 136 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 138 138 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 142 142 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 152 152 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 163 163 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 166 166 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 167 167 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 168 168 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 171 171 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 173 173 CALCIUM 1 AND 3 (BY SIMILARITY).  
 FT SEQUENCE 438 AA; 47796 MW; 2393FDAC51B80F51 CRC64;  
 Query Match 17.9%; Score 365.5; DB 1; Length 438;  
 Best Local Similarity 24.2%; Pred. No. 4.9e-19;  
 Matches 109; Conservative 64; Mismatches 132; Indels 95; Gaps 18;  
 QY 12 VIMQAFWD--VPSGGIWDITROKIPENTYDAGISAIWTPASKMGAGYSGYDPPDF 69  
 DB 27 VLFQGNWESKQSGG-WYNNMMKVDVIAAGVTHVLPSPSHSVS---NEGYPGRLY 82  
 QY 70 DLGEYDQKGTVEFRFGKQELVNMINTAHAYGMKVIADIVINHR-----113  
 DB 83 DID-----ASYGNAELKSLIGALHGKGVQAIADIVINHRCAADYKDSRGIVCEP 134  
 QY 114 --AGDLEWRFVNDYTWTPSKVASKYATNYLDFHFNELHAGDSGT-FGGYDICH--168  
 DB 135 GTSGLRDWGHM--ICRDTKYSDG--TANL-----DTGADFAAAPDIDHLN 178  
 QY 169 ---DKSWDQYWLWASQESYAAYLRSIGIDAWRFDYKGYAPVWVVDKLNWNG-GNAVGEY 224  
 DB 179 DRVQRELKELLWLKSD-----LGFDAWRLDFARGYSPEWAKYVIDGTSPSLAVAEV 230  
 QY 225 WDT-----NVDA-----VLNW-----AYSSGAKVDFPALYKMDAEPDNKNIPAL 264  
 DB 231 WDNMATGSGDKPNYDODAHQNLVNWVDRKVGGAASAGWVDFTTKGLNAAVEGE-LWRL 289  
 QY 265 VSALQNGQTVVSRDPFXAVTFVANHDTDIWNKYP-----AYAFILTYEGQPTFYR 316  
 DB 290 IDPQKAPGVGWGPAPAKAIFVNDHDTGSTQAWMPFSDKWQGYAYILTHPGICIFVD 349  
 QY 317 DYEWLNKDKLNLIWHEN---LAGGSTDIYVDNDELI 353  
 DB 350 HFFWNGFQDQIALVAIRKNGITATSALKILMHGDAYV 389  
 RESULT 12  
 ID AMY3 HORVU  
 AC P04747; STANDARD; PRT; 368 AA.  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-  
 DE glucan glucanohydrolase) (Clone PHV19) (Fragment).  
 GN AMY1.3.  
 OS Hordeum vulgare (Barley).

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
 Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chandler P.M., Zwar J.A., Jacobsen J.V., Higgins T.J.V., Inglis A.S.;  
 RT "The effects of gibberellic acid and abscisic acid on alpha-amylase  
 RT mRNA levels in barley aleurone layers studies using an alpha amylase  
 RT cDNA clone."; Plant Mol. Biol. 3:407-418(1984).  
 RL -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally  
 CC regulated. Germinating embryos produce the hormone gibberellic  
 CC acid, which within 10 hours stimulates the aleurone cells covering  
 CC the endosperm of the seed to produce alpha-amylase. The enzyme  
 CC then degrades the starch within the endosperm for use by the  
 CC developing plant embryo.  
 CC -1- MISCELLANEOUS: There are at least 4 types of alpha-amylase in  
 CC barley.  
 CC -1- MISCELLANEOUS: Type B isozyme mRNA is undetectable in unstimulated  
 CC cells and increases a hundred-fold after stimulation with  
 CC gibberellic acid.  
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 DR EMBL; K02638; AAA32933.1; -.  
 DR HSP; P04063; IAVA.  
 DR InterPro: IPR006589; Alp amyl cat sub.  
 DR InterPro: IPR006047; Alpha amyl cat.  
 DR InterPro: IPR006046; Glyco\_hydro\_13.  
 DR Pfam: PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SMO0642; Amy; 1.  
 KW Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
 KW Calcium-binding; Signal; Multigene family.  
 FT SIGNAL 1 24  
 FT CHAIN 25 368 ALPHA-AMYLASE TYPE B ISOZYME.  
 FT ACT\_SITE 203 203 BY SIMILARITY.  
 FT ACT\_SITE 228 228 BY SIMILARITY.  
 FT ACT\_SITE 313 313 BY SIMILARITY.  
 FT METAL 115 115 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 132 132 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 135 135 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 137 137 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 141 141 CALCIUM 2 (BY SIMILARITY).  
 FT METAL 151 151 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 162 162 CALCIUM 3 (BY SIMILARITY).  
 FT METAL 165 165 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 166 166 CALCIUM 1 (BY SIMILARITY).  
 FT METAL 167 167 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 170 170 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY SIMILARITY).  
 FT METAL 172 172 CALCIUM 1 AND 3 (BY SIMILARITY).  
 FT NON\_TER 368 368  
 FT SEQUENCE 368 AA; 40787 MW; A237EF55793BA93B CRC64;  
 Query Match 14.8%; Score 363.5; DB 1; Length 368;  
 Best Local Similarity 27.0%; Pred. No. 5.6e-19;  
 Matches 106; Conservative 55; Mismatches 118; Indels 113; Gaps 16;





DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Alpha-amylase isozyme 3C precursor (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE glucanohydrolase).  
 GN AMY1.7 OR AMY3B.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OC NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Japonica M202; TISSUE=etiolated leaf;  
 RX MEDLINE=91329692; PubMed=1714318;  
 RA Sutliff T.D., Huang N., Litts J.C., Rodriguez R.L.;  
 RT "Characterization of an alpha-amylase multigene cluster in rice.";  
 RL Plant Mol. Biol. 16:579-591(1991).  
 CC !- FUNCTION: Important for breakdown of endosperm starch during  
 CC germination.  
 CC !- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC !- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC !- SUBUNIT: Monomer.  
 CC !- TISSUE SPECIFICITY: Germinating seeds.  
 CC !- DEVELOPMENTAL STAGE: Expressed at a high level during germination  
 CC in the aleurones cells under the control of the plant hormone  
 CC gibberellic acid and in the developing grains at a low level.  
 CC !- SIMILARITY: Belongs to family 13 of Glycosyl hydrolases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; X56338; CAA39778.1; --  
 DR PIR; S14956; S14956.  
 DR HSSP; P04063; 1A.VA.  
 DR Gramene; P27939; --  
 DR InterPro; IPR006589; Alp\_aml\_cat\_sub.  
 DR InterPro; IPR006047; Alpha\_aml\_cat.  
 DR InterPro; IPR006046; Glyco\_Hydro\_13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SM00642; Aamy; 1.  
 DR Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 26  
 FT CHAIN 27 437  
 FT ACT SITE 205 205  
 FT ACT SITE 313 313  
 FT ACT SITE 313 313  
 FT METAL 117 117  
 FT METAL 134 134  
 FT METAL 137 137  
 FT METAL 139 139  
 FT METAL 143 143  
 FT METAL 153 153  
 FT METAL 164 164  
 FT METAL 167 167  
 FT METAL 168 168  
 FT METAL 169 169  
 FT METAL 172 172  
 FT METAL 174 174  
 SQ SEQUENCE 437 AA; 48637 MW; BD304250B40C7A8B CRC64;  
 Query Match 14.4%; Score 353; DB 1; Length 437;  
 Best Local Similarity 26.4%; Pred. No. 3.9e-18;

Matches 108; Conservative 55; Mismatches 138; Indels 108; Gaps 18;  
 QY 5 SELEKGVIMQAFYNDV--PSGGIWWDTIRQKIPWYDAGISAIWIPASKCMGAYSMG 62  
 DB 21 SHLAQAQVLFQGFNWSNKKGG-WYNFLHSHVDYIAATGVTHVWLPSPSHSVA---PQG 76  
 QY 63 YDPYDFDLGEYDQKGTVEFRFGSKGLVNMINTAHAYGMKVIADIVINHRA----- 114  
 DB 77 YMPGRLYDLD-----ASKYGTGAELSLIAAHPHSKSIKCAVDIVINHRCADYKDSRG 128  
 QY 115 -----GGD-----LEWNPFFVNDYTWDFSKVASKYKTYANLDFHFNELHAGDSGTGFGYP 164  
 DB 129 IYCFEGGTPDSRLDWGP---DMICSDDTQYNSRG-----HRDTGADFGAAP 173  
 QY 165 DICHKDSWDQY---WL-WASQESYAAVLRISIGIDAWREDYVKGYPWVVDWL-NMGG 218  
 DB 174 DIDHLNTRVQTELSDLNMLKSD-----VGFDGWRLLDPAKGYSAIVAKTYVDNTDPS 225  
 QY 219 WYVGEYH-----DTNVDVAVNWAYSSG--AKVPDFALYYKXDEAFD---- 257  
 DB 226 FVVAEIMSNMRYDNGEPPSNQDGRQELVNAQAVGGPASFDTTKGELQAAVQGSWM 285  
 QY 258 -----NKNIPALVSALQNGQTVVRDPFRAVTFVANHDTDIWNKYP-----AYAFI 304  
 DB 286 RMKDGNGKAPGMIGWL-----PEKAVTFIDNHDGTGTSQNSWPFPSDKVMQRYAVI 335  
 QY 305 LTVEGQPTFYRDEYEWLNKDKLNIWHENLA---GGSTDIVYYDND 350  
 DB 336 LTHPGVPCFYDHVFNQNLKQEIETLAAVRSRNGIHGSKNLILAAADGD 384  
 RESULT 15  
 ID AMY6 HORVU STANDARD; PRT; 429 AA.  
 AC P04750;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-  
 DE glucan glucanohydrolase) (Clones GRAMy56 and 963).  
 GN AMY1.6  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
 OC Triticeae; Hordeum.  
 OC NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A. (CLONE GRAMy56).  
 RA Rahmatullah R.J., Huang J.-X., Clark K.L., Reeck G.R.,  
 RA Chandra G.R., Muthukrishnan S.;  
 RT "Nucleotide and predicted amino acid sequences of two different genes  
 RT for high-pi alpha-amylases from barley.";  
 RL Plant Mol. Biol. 12:119-121(1989).  
 RN [2]  
 RP SEQUENCE OF 380-429 FROM N.A. (CLONE 963).  
 RX MEDLINE=85159405; PubMed=6335720;  
 RA Huang J.-X., Swegle M., Dandekar A.M., Muthukrishnan S.;  
 RT "Expression and regulation of alpha-amylase gene family in barley  
 RT aleurones.";  
 RL J. Mol. Appl. Genet. 2:579-588(1984).  
 CC !- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC !- COFACTOR: Binds 3 calcium ions per subunit (By similarity).  
 CC !- SUBUNIT: Monomer (By similarity).  
 CC !- DEVELOPMENTAL STAGE: Production of alpha-amylase is hormonally  
 CC regulated. Germinating embryos produce the hormone gibberellic  
 CC acid, which within 10 hours stimulates the aleurone cells covering  
 CC the endosperm of the seed to produce alpha-amylase. The enzyme  
 CC then degrades the starch within the endosperm for use by the  
 CC developing plant embryo.  
 CC !- MISCELLANEOUS: There are at least 4 types of alpha-amylase in  
 CC barley.  
 CC !- MISCELLANEOUS: Type B isozyme mRNA is undetectable in unstimulated

CC cells and increases a hundred-fold after stimulation with  
CC gibberellic acid.  
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X15227; CAA33299.1; -;  
CC EMBL; K02636; AAA32932.1; -;  
CC PIR; JEO406; JEO406.  
CC HSP; P04063; IAVA.  
CC InterPro; IPR006589; Alp.amyl.cat.sub.  
CC InterPro; IPR006047; Alpha.amyl.cat.  
CC InterPro; IPR006046; Glyco\_hydro\_13.  
CC Pfam; PF00128; alpha-amylase; 1.  
CC PRINTS; PR00110; ALPHAAMYLASE.  
CC SMART; SM00642; Amy; 1.  
KW Germination; Carbohydrate metabolism; Hydrolase; Glycosidase; Seed;  
KW Calcium-binding; Signal; Multigene family.  
FT SIGNAL 1 24  
FT CHAIN 25 429 ALPHA-AMYLASE TYPE B ISOZYME.  
FT ACT\_SITE 205 205 BY SIMILARITY.  
FT ACT\_SITE 315 315 BY SIMILARITY.  
FT METAL 115 115  
FT METAL 132 132 CALCIUM 1 (BY SIMILARITY).  
FT METAL 135 135 CALCIUM 2 (BY SIMILARITY).  
FT METAL 137 137 CALCIUM 2 (BY SIMILARITY).  
FT METAL 141 141 CALCIUM 2 (BY SIMILARITY).  
FT METAL 151 151 CALCIUM 3 (BY SIMILARITY).  
FT METAL 167 167 CALCIUM 3 (VIA CARBONYL OXYGEN) (BY  
FT METAL 168 168 SIMILARITY).  
FT METAL 169 169 CALCIUM 1 (BY SIMILARITY).  
FT METAL 172 172 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY  
FT METAL 174 174 SIMILARITY).  
FT METAL 174 174 CALCIUM 1 AND 3 (BY SIMILARITY).  
FT SEQUENCE 429 AA; 47937 MW; 1C924CA6319D5262 CRC64;  
Query Match 14.3%; Score 351.5; DB 1; Length 429;  
Best Local Similarity 27.1%; Pred. No. 4.8e-18;  
Matches 106; Conservative 54; Mismatches 122; Indels 109; Gaps 17;  
QY 7 LEKGVIVQAFYWDV-PSGGIWDITROKIPEDWDAGISAIWIPPASKMGGMGYDP 65  
DB 21 LASGVLFQGFENWESKNGWYFNLMGKVDDIAAGVTHVWLPASQSA---EQGYMP 77  
QY 66 YDFFDLGEYDQGTETRFSGKQELVNMINTAHAYGMKVIADIVNHRA----- 114  
DB 78 GELYLD-----ASKYGNKAQLKSLIGALHGKAVKAIADIVINHRTAERKDGRIYC 129  
QY 115 ---GSD---LEWNPV---NDYTWDFSKVASKYTYANYLDHPNHLHAGDSGTFGY 164  
DB 130 IFEGGTPDARLDWGHMICRDRPYPD-----GTGN---RPTRTA---DFGAAP 173  
QY 165 DICH-----DKSDQYWLWASQESYAAYLRISIGIDAKRFDYVGYAPWVKDWLN-WWGG 218  
DB 174 DIDHLNPRVQKELVEWLNLRITDD-----GFGWRFDFAKGSADVAKIYYDRSEPS 225  
QY 219 WAVEGYW-----DTNVDVNLWAYSSG-----AKVDFPALYYKMDFAFDNK 259  
DB 226 FAVAEIWTSLAYGGDGKFNQDQPHRQELVNMVNVKVGSGPATTFDTTKGILNVAVEGE 285  
QY 260 -----NIPALVSALONGQTVVSRDPFKAVTFVANHDTIWNKYP-----AYA 302  
DB 286 LWRRLGTDGKAPGMIGW-----PAKAVTFVDNHDGTGQCHWMPFSDRWMOGYA 335  
QY 303 FILTYEGQPTFYRDYEEWLNKDKLNLIWI 333

Db 336 YILTHPGNPCIIFYDHFDFMGLKEEDRLVSI 366  
Search completed: June 29, 2004, 09:34:44  
Job time : 19 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:31:27 ; Search time 45 Seconds  
(without alignments)  
3057.021 Million cell updates/sec

Title: US-10-081-739A-2

Perfect score: 2459

Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQYGVNSYGVGV 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_muc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2311	94.0	432	14 Q8JZK3	Q8JZK3 uncultured
2	2298	93.5	461	1 Q8NKR5	Q8NKR5 thermococcus
3	2279	92.7	461	1 Q8NKR4	Q8NKR4 thermococcus
4	2248	91.4	461	1 Q33476	Q33476 pyrococcus
5	2221	90.3	457	1 Q93647	Q93647 thermococcus
6	2199.5	89.4	469	1 Q50200	Q50200 thermococcus
7	2153	87.6	460	1 Q84522	Q84522 pyrococcus
8	2153	87.6	473	17 Q8U319	Q8U319 pyrococcus
9	2152	87.5	460	1 Q9P910	Q9P910 pyrococcus
10	513	20.9	519	2 Q9Q788	Q9Q788 cytophaga s
11	511.5	20.8	507	16 Q87HG6	Q87HG6 vibrio para
12	492	20.0	492	16 Q8YUZ1	Q8YUZ1 anabaena sp
13	487.5	19.8	494	16 Q8Z5S5	Q8Z5S5 salmonella
14	486	19.8	501	2 Q31148	Q31148 bacillus sp
15	472.5	19.2	516	2 Q82839	Q82839 bacillus sp
16	469.5	19.1	549	2 Q31193	Q31193 bacillus st

17	468.5	19.1	549	2	Q9KXV6	Q9KXV6 bacillus st
18	467.5	19.0	521	16	Q8IAS4	Q8IAS4 bacillus ce
19	464.5	18.9	531	2	P71034	P71034 bacillus sp
20	464.5	18.9	533	2	Q9AQ54	Q9AQ54 bacillus me
21	456	18.5	613	2	Q59222	Q59222 bacillus sp
22	455.5	18.5	513	16	Q81YJ4	Q81YJ4 bacillus an
23	452.5	18.4	495	16	Q8XBB6	Q8XBB6 escherichia
24	451.5	18.4	495	16	Q8JRA0	Q8JRA0 shigella fl
25	451.5	18.4	506	16	Q8U916	Q8U916 agrobacteri
26	448.5	18.2	495	16	Q8FGL8	Q8FGL8 escherichia
27	446.5	18.2	495	16	Q7UAB0	Q7UAB0 shigella fl
28	438	17.8	491	16	Q9CG59	Q9CG59 lactococcus
29	435	17.7	481	16	Q89YF1	Q89YF1 bacteroides
30	420	17.1	493	2	Q03657	Q03657 bacillus ci
31	412.5	16.8	420	10	Q9ZP43	Q9ZP43 phaseolus v
32	406	16.5	421	10	Q7X9T1	Q7X9T1 phaseolus a
33	398	16.2	423	10	Q42678	Q42678 cuscuta ref
34	392	15.9	424	10	Q8LP27	Q8LP27 pharbitis n
35	385.5	15.7	488	16	Q8E696	Q8E696 streptococ
36	385.5	15.7	488	16	Q8E0M2	Q8E0M2 streptococ
37	379.5	15.4	484	2	O50583	O50583 streptococ
38	376.5	15.3	421	10	Q42504	Q42504 hordeum vul
39	373.5	15.2	437	10	O4965	O4965 hordeum vul
40	373	15.2	484	16	Q97Q49	Q97Q49 streptococ
41	372	15.1	484	16	Q8DPC8	Q8DPC8 streptococ
42	371	15.1	403	10	Q7Y1C3	Q7Y1C3 eleusine co
43	365.5	14.9	407	10	O41442	O41442 solanum tub
44	363.5	14.8	504	1	Q60224	Q60224 natronococ
45	362	14.7	416	10	Q8LJQ6	Q8LJQ6 musa acumin

#### ALIGNMENTS

#### RESULT 1

Q8JZK3 PRELIMINARY; PRT; 432 AA.  
ID Q8JZK3  
AC Q8JZK3  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Alpha-amylase.  
GN B5063.  
OS uncultured organism.  
OC unclassified; environmental samples.  
OX NCBI\_TaxID=155900;  
RN [1]  
RP MEDLINE-22113025; PubMed-11994309;  
RX Richardson T.H., Fan X., Frey G., Callen W., Cabell M., Lam D.,  
RA Macomber J., Short J.M., Robertson D.E., Miller C.;  
RT "A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE.";  
J. Biol. Chem. 277:26501-26507(2002).  
RL EMBL: AF504064; AAM48114.1; -;  
DR GO: GO:0004556; P:alpha-amylase activity; IEA.  
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro: IPR006047; Alpha\_amyl\_cat.  
DR InterPro: IPR006589; Alp\_amyl\_cat sub.  
DR Pfam: PF00128; Glyco\_hydro\_I3.  
DR PRINTS: PR00110; ALPHAAMYLASE.  
DR SMART: SM00642; Amy; 1.  
SQ SEQUENCE 432 AA; 49225 MW; ECE4A87826D31D2D CRC64;

Query Match 94.0%; Score 2311; DB 14; Length 432;  
Best Local Similarity 94.7%; Pred. No. 3e-148;  
Matches 407; Conservative 10; Mismatches 13; Indels 0; Gaps 0;  
QY 7 LEKGGVIMQAFYWDVPSGGIWDITROKIPEDYDAGISAIWIIPDASKMGAGYSGYDYP 66  
DB 3 LEEGGLIMQAFYWDVPMGGIWDITROKIPEDYDAGISAIWIIPDASKMGAGYSGYDYP 62

QY 67 DRPDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVINHRAGGDLWNPFFVND 126  
DB 63 DYFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVINHRAGGDLWNPFFVND 122  
QY 127 YTWDFSKVASGKYTANYLDLDFHNPHELHAGDSGTFGGYPDI CHDKSWDQYWLWASQESVAA 186  
DB 123 YTWDFSKVASGKYTANYLDLDFHNPHELHAGDSGTFGGYPDI CHDKSWDQYWLWASQESVAA 182  
QY 187 YLRSIGIDAWRFDYVKGYPVYKQWLVNMGWAGVGEYDNTVDVAVLNWYSSGAKVDFD 246  
DB 183 YLRSIGIDAWRFDYVKGYPVYKQWLVNMGWAGVGEYDNTVDVAVLNWYSSGAKVDFD 242  
QY 247 ALYKQDEAFDNKNI PALYSALQNGQTVVSRDPFKAVTFVANHDTDI IWNKYPAYAFILT 306  
DB 243 ALYKQDEAFDNKNI PALVDALRYGQTVVSRDPFKAVTFVANHDTDI IWNKYPAYAFILT 302  
QY 307 YEQOPTIFRYDYEELNKKLNIWIHNHLAGGSTDI VYDNDDELIFVRNGYGD 366  
DB 303 YEQOPTIFRYDYEELNKKLNIWIHNHLAGGSTDI VYDNDDELIFVRNGYGD 362  
QY 367 TYINLGSSKAGRWVYVYKPFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPANGQY 426  
DB 363 TYINLGSSKAGRWVYVYKPFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPANGQY 422  
QY 427 YSVWSYCGVG 436  
DB 423 YSVWSYCGVG 432

RESULT 2  
Q8NKR5 PRELIMINARY; PRT; 461 AA.  
AC Q8NKR5; 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Alpha-amylase precursor.  
GN BD5064.  
OS Thermococcus sp. GUSL5.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Thermococcus.  
OX NCBI\_TaxID=195100;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GUSL5;  
RX MEDLINE=22113025; PubMed=11994309;  
RA Richardson T.H., Tan X., Frey G., Callen W., Cabell M., Lam D.,  
RA Macomber J., Short J.M., Robertson D.E., Miller C.;  
RT "A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE.";  
RL J. Biol. Chem. 277:26501-26507(2002).  
DR EMBL; AF504062; AAM48112.1; -;  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR InterPro; IPR006589; Alp amyl cat sub.  
DR InterPro; IPR006046; Glyco hydro I3.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PRINTS; PR00128; alpha-amylase; 1.  
DR SMART; SM00642; Amy; 1.  
FT SIGNAL.  
KW SIGNAL.  
SQ SEQUENCE 461 AA; 52328 MW; 57F0DD3AB58EDACB CRC64;  
Query Match 93.5%; Score 2298; DB 1; Length 461;  
Best Local Similarity 93.1%; Pred. No. 2.5e-147;  
Matches 405; Conservative 14; Mismatches 16; Indels 0; Gaps 0;  
QY 2 AKYSELEKGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGWGAYSM 61  
DB 27 AKYLEEGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGWGAYSM 86  
QY 62 GYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVINHRAGGDLWNP 121

DB 87 GYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVINHRAGGDLWNP 146  
QY 122 PFVNDYTWDFSKVASGKYTANYLDLDFHNPHELHAGDSGTFGGYPDI CHDKSWDQYWLWASQ 181  
DB 147 PFVNDYTWDFSKVASGKYTANYLDLDFHNPHELHAGDSGTFGGYPDI CHDKSWDQYWLWASQ 206  
QY 182 ESAAALYRSIGIDAWRFDYVKGYPVYKQWLVNMGWAGVGEYDNTVDVAVLNWYSSGA 241  
DB 207 ESAAALYRSIGIDAWRFDYVKGYPVYKQWLVNMGWAGVGEYDNTVDVAVLNWYSSGA 266  
QY 242 KYDFALYKQDEAFDNKNI PALYSALQNGQTVVSRDPFKAVTFVANHDTDI IWNKYPAY 301  
DB 267 KYDFALYKQDEAFDNKNI PALVDALRYGQTVVSRDPFKAVTFVANHDTDI IWNKYPAY 326  
QY 302 AFILTYEGQPTIFRYDYEELNKKLNIWIHNHLAGGSTDI VYDNDDELIFVRNGYGD 361  
DB 327 AFILTYEGQPTIFRYDYEELNKKLNIWIHNHLAGGSTDI VYDNDDELIFVRNGYGD 386  
QY 362 KPGIITINLGSSKAGRWVYVYKPFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPA 421  
DB 387 KPGIITINLGSSKAGRWVYVYKPFAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPA 446  
QY 422 NGQYGSVWSYCGVG 436  
DB 447 NGYGSVWSYCGVG 461  
RESULT 3  
Q8NKR4 PRELIMINARY; PRT; 461 AA.  
AC Q8NKR4; 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Alpha-amylase precursor.  
GN BD5031.  
OS Thermococcus sp. 'AEP11 1a'.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Thermococcus.  
OX NCBI\_TaxID=195101;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AEP11 1a;  
RX MEDLINE=22113025; PubMed=11994309;  
RA Richardson T.H., Tan X., Frey G., Callen W., Cabell M., Lam D.,  
RA Macomber J., Short J.M., Robertson D.E., Miller C.;  
RT "A Novel, High Performance Enzyme for Starch Liquefaction. DISCOVERY  
AND OPTIMIZATION OF A LOW pH, THERMOSTABLE alpha -AMYLASE.";  
RL J. Biol. Chem. 277:26501-26507(2002).  
DR EMBL; AF504063; AAM48113.1; -;  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR InterPro; IPR006589; Alp amyl cat sub.  
DR InterPro; IPR006046; Glyco hydro I3.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PRINTS; PR00110; ALPHAAMYLASE.  
DR SMART; SM00642; Amy; 1.  
FT SIGNAL.  
KW SIGNAL.  
SQ SEQUENCE 461 AA; 52378 MW; 17505F28DE3F6953 CRC64;  
Query Match 92.7%; Score 2279; DB 1; Length 461;  
Best Local Similarity 91.5%; Pred. No. 4.8e-146;  
Matches 398; Conservative 17; Mismatches 20; Indels 0; Gaps 0;  
QY 2 AKYSELEKGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGWGAYSM 61  
DB 27 AKYLEEGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGWGAYSM 86  
QY 62 GYDPYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVADIIVINHRAGGDLWNP 121

Db 87 GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGKIADIVINHRAGDLEWN 146  
QY 122 PFVNDYTWTFDFSKVASGKYTANYLDFHPNHLHAGDSGTFGGYDPICHDKSWDQYWLWASQ 181  
Db 147 PFVGDYTWTFDFSKVASGKYTANYLDFHPNEVNEKCCDEGTFGGPDIAHEKSWDQHWLWASD 206  
QY 182 ESYAAYLRSIGIDAWRFDVVKGYPVWVKDLNWNWGGWAGVYDWTNDVAVLNWYSSGA 241  
Db 207 ESYAAYLRSIGIDAWRFDVVKGYPVWVKDLNWNWGGWAGVYDWTNDVAVLNWYSSGA 266  
QY 242 KVDFDALYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPAY 301  
Db 267 KVDFDALYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPAY 326  
QY 302 APILTYEGOPTIFRYDYEELWLNKDLNLIWHENLAGGSTDIYVYDNDLIFVRNGYGD 361  
Db 327 APILTYEGOPTIFRYDYEELWLNKDLNLIWHENLAGGSTDIYVYDNDLIFVRNGYGD 386  
QY 362 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPA 421  
Db 387 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPA 446  
QY 422 NGQYGYVSWSYCGVG 436  
Db 447 NGQYGYVSWSYCGVG 461

RESULT 4  
C33476 PRELIMINARY; PRT; 461 AA.  
ID O33476  
AC O33476  
DT 01-JAN-1998 (TremBLrel. 05, Created)  
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)  
DE Alpha-amylase precursor.  
GN APK.  
OS Pyrococcus kodakaraensis.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Thermococcus.  
OX NCBI\_TaxID=69014;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KOD1;  
RA Tachibana Y., Mendez L., Fujiwara S., Takagi M., Imanaka T.;  
RT "Cloning and expression of the alpha-amylase gene from the  
RT hyperthermophilic archaeon Pyrococcus sp. KOD1, and characterization  
RT of the enzyme."  
RL J. Ferment. Bioeng. 82:224-232(1996).  
DR EMBL; D83793; BAR21130.1; -.  
DR HSSP; P06278; IVUS.  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR Pfam; PF00128; Glyco\_hydro\_13.  
DR PRINTS; PR00110; ALPHAAMYLASE.  
KW Signal.  
FT SIGNAL  
FT CHAIN  
SQ SEQUENCE 461 AA; 52213 MW; FCC131A93DC03123 CRC64;

Query Match 91.4%; Score 2248; DB 1; Length 461;  
Best Local Similarity 90.6%; Pred. No. 5.9e-144;  
Matches 394; Conservative 18; Mismatches 23; Indels 0; Gaps 0;

QY 2 AKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPPSKMGGGAYSM 61  
Db 27 AKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPPSKMGGGAYSM 86  
QY 62 GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGKIADIVINHRAGDLEWN 121  
Db 87 GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGKIADIVINHRAGDLEWN 146

QY 122 PFVNDYTWTFDFSKVASGKYTANYLDFHPNHLHAGDSGTFGGYDPICHDKSWDQYWLWASQ 181  
Db 147 PFVGDYTWTFDFSKVASGKYTANYLDFHPNEVNEKCCDEGTFGGPDIAHEKSWDQHWLWASD 206  
QY 182 ESYAAYLRSIGIDAWRFDVVKGYPVWVKDLNWNWGGWAGVYDWTNDVAVLNWYSSGA 241  
Db 207 ESYAAYLRSIGIDAWRFDVVKGYPVWVKDLNWNWGGWAGVYDWTNDVAVLNWYSSGA 266  
QY 242 KVDFDALYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPAY 301  
Db 267 KVDFDALYKMDAEAFDNKNI PALVSALQNGQTVVSRDPFKAVTFVANHTDIIWNKYPAY 326  
QY 302 APILTYEGOPTIFRYDYEELWLNKDLNLIWHENLAGGSTDIYVYDNDLIFVRNGYGD 361  
Db 327 APILTYEGOPTIFRYDYEELWLNKDLNLIWHENLAGGSTDIYVYDNDLIFVRNGYGD 386  
QY 362 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPA 421  
Db 387 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKYVYSGWVYLEAPAYDPA 446  
QY 422 NGQYGYVSWSYCGVG 436  
Db 447 NGQYGYVSWSYCGVG 461

RESULT 5  
O93647 PRELIMINARY; PRT; 457 AA.  
ID O93647  
AC O93647  
DT 01-MAY-1999 (TremBLrel. 10, Created)  
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)  
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)  
DE Alpha-amylase (EC 3.2.1.1).  
GN AMY.  
OS Thermococcus hydrothermalis.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Thermococcus.  
OX NCBI\_TaxID=46539;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AL662;  
RX MEDLINE=20243462; PubMed=10779714;  
RA Leveque E., Haye B., Belarbi A.;  
RT "Cloning and expression of an alpha-amylase encoding gene from the  
RT hyperthermophilic archaeobacterium Thermococcus hydrothermalis and  
RT biochemical characterization of the recombinant enzyme."  
RL FEMS Microbiol. Lett. 186:67-71(2000).  
DR EMBL; AF068255; AAC97877.1; -.  
DR HSSP; P06278; IVUS.  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR Pfam; PF00128; alpha-amylase; i.  
KW Glycosidase; Hydrolase.  
SQ SEQUENCE 457 AA; 51543 MW; 602367643AC7847B CRC64;

Query Match 90.3%; Score 2221; DB 1; Length 457;  
Best Local Similarity 89.9%; Pred. No. 3.9e-142;  
Matches 391; Conservative 20; Mismatches 24; Indels 0; Gaps 0;

QY 2 AKYSELEKGGVIMQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPPSKMGGGAYSM 61  
Db 23 AKATLENGGVIMQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPPSKMGGGAYSM 82  
QY 62 GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGKIADIVINHRAGDLEWN 121  
Db 83 GYDPYDFDLGEYNQKGTIETRFSGKQELINNTAHAYGKIADIVINHRAGDLEWN 142  
QY 122 PFVNDYTWTFDFSKVASGKYTANYLDFHPNHLHAGDSGTFGGYDPICHDKSWDQYWLWASQ 181  
Db 143 PFTNSYTWTFDFSKVASGKYTANYLDFHPNHLHAGDSGTFGGYDPICHDKSWDQHWLWASN 202

QY 182 ESVAALRSIGIDAMRFDYKGYAPVWVWDLNMGWNAVGYBDTNDVAVLNWAYSSGA 241  
 Db 203 ESVAALRSIGIDAMRFDYKGYAPVWVWDLNMGWNAVGYBDTNDVAVLNWAYSSGA 262  
 QY 242 KVFDFALYKMDAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPAY 301  
 Db 263 KVFDFALYKMDAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPAY 322  
 QY 302 AFILTYEGOPTFYRDEEWNKDKLNIWIHENLAGGSTDIVVYDNDLIFVRNGYDXPGLI 361  
 Db 323 AFILTYEGOPTFYRDEEWNKDKLNIWIHENLAGGSTDIVVYDNDLIFVRNGYDXPGLI 382  
 QY 362 KPLIITYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 421  
 Db 383 KPLIITYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 442  
 QY 422 NGQGYVMSYCGVG 436  
 Db 443 NGQGYVMSYCGVG 457  
 RESULT 6  
 ID 050200 PRELIMINARY; PRT; 469 AA.  
 AC 050200;  
 DT 01-JUN-1998 (T=EMBLrel. 06, Created)  
 DT 01-JUN-1998 (T=EMBLrel. 06, Last sequence update)  
 DT 01-JUN-2003 (T=EMBLrel. 24, Last annotation update)  
 DE Amylase (EC 3.2.1.1).  
 GN AMY.  
 OS Thermococcus sp. Rt3.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Thermococcus.  
 OX NCBI\_TaxID=65421;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RT3;  
 RX MEDLINE=99154282; PubMed=10030014;  
 RA Jones R.A., Jermin L.S., Eastal S., Patel B.K., Beacham I.R.;  
 RT "Amylase and 16S rRNA genes from a hyperthermophilic  
 archaeobacterium."  
 RL J. Appl. Microbiol. 86:93-107(1999).  
 DR EMBL; AF017454; AAB87860.1; -.  
 DR HSP; P06278; 1VJS.  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha-amyl cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 469 AA; 52756 MW; C3BD614FBA4DB3E2 CRC64;  
 Query Match 89.4%; Score 2199.5; DB 1; Length 469;  
 Best Local Similarity 90.5%; Pred. No. 1.1e-140;  
 Matches 389; Conservative 16; Mismatches 24; Indels 1; Gaps 1;  
 QY 7 LEKGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPPASKMGAYSMGYDPP 66  
 Db 41 LEBGGVIMQAFYWDVPSGGIWDITRQKIPWYDAGISAIWIPPPASKMGAYSMGYDPP 100  
 QY 67 DFDLGEYDQKVTETRFSGKQLVNMINTAHAYGMKVADIVINHRAGDLEWPFVND 126  
 Db 101 DFDLGEYDQKVTETRFSGKQLVNMINTAHAYGMKVADIVINHRAGDLEWPFVND 160  
 QY 127 YTWDSKVASGKYTANYLDHFNELHAGDSGTFGGYDIPICDKSWDQYWLWASQESVAA 186  
 Db 161 YTWDSKVASGKYTANYLDHFNELHAGDSGTFGGYDIPICDKSWDQYWLWASQESVAA 220  
 QY 187 YLRSIGIDAMRFDYKGYAPVWVWDLNMGWNAVGYBDTNDVAVLNWAYSSGA 246  
 Db 221 YLRSIGIDAMRFDYKGYAPVWVWDLNMGWNAVGYBDTNDVAVLNWAYSSGA 279

QY 247 ALYKMDAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPAYAFILT 306  
 Db 280 ALYKMDAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHTDIIWNKYPAYAFILT 339  
 QY 307 YEGOPTFYRDEEWNKDKLNIWIHENLAGGSTDIVVYDNDLIFVRNGYDXPGLI 366  
 Db 340 YEGOPTFYRDEEWNKDKLNIWIHENLAGGSTDIVVYDNDLIFVRNGYDXPGLI 399  
 QY 367 TYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 426  
 Db 400 TYNLGSSKAGRWYVPKFAACIHEYTGNLGGWVDKYVSSGWVYLEAPADPA 459  
 QY 427 YSVMSYCGVG 436  
 Db 460 YSVMSYCGVG 469  
 RESULT 7  
 ID 008452 PRELIMINARY; PRT; 460 AA.  
 AC 008452;  
 DT 01-JUL-1997 (T=EMBLrel. 04, Created)  
 DT 01-JUL-1997 (T=EMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (T=EMBLrel. 25, Last annotation update)  
 DE Alpha amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase)  
 DE (GLYCOGENASE).  
 GN AMYA.  
 OS Pyrococcus furiosus, and  
 OS Pyrococcus woesei.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=2261, 2262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.furiosus; STRAIN=DSM 3638;  
 RA Dong G., Vieille C., Savchenko A., Zeikus J.G.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.furiosus; STRAIN=DSM3638;  
 RX MEDLINE=97341170; PubMed=9195939;  
 RA Jorgensen S., Vorgias C.E., Antranikian G.;  
 RT "Cloning, sequencing, characterization, and expression of an  
 extracellular alpha-amylase from the hyperthermophilic archaeon  
 Pyrococcus furiosus in Escherichia coli and Bacillus subtilis."  
 RL J. Biol. Chem. 272:16335-16342(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.furiosus; STRAIN=DSM3638;  
 RA Jorgensen S.T., Vorgias C.E., Antranikian G.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=P.woesei; STRAIN=DSM3773;  
 RA Chunlin Lu, Weizheng J., Yunyan Y.;  
 RT "Cloning and Expression of Alpha Amylase from Pyrococcus woesei."  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 DR EMBL; AF001268; AAC45663.1; -.  
 DR EMBL; U96622; AAB67705.1; -.  
 DR EMBL; AF177906; AAD54338.1; -.  
 DR HSP; P06278; 1VJS.  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha amyl cat.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 460 AA; 52909 MW; 0E0A2AEB5FCAF541 CRC64;



Query Match 87.6%; Score 2153; DB 1; Length 460;  
Best Local Similarity 86.7%; Pred. No. 1.6e-137;  
Matches 377; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

2 AKYLEEGGVIMQAFYWDVPSGGIWDTRIROPKIPWYDAGISAIWIPPSKMGSGGYSM 61  
26 AKYLEEGGVIMQAFYWDVPSGGIWDTRIROPKIPWYDAGISAIWIPPSKMGSGGYSM 85  
62 GYDPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEWN 121  
86 GYDPDYDFDLGEYDQKGTETRFSGKQELVRLIQTAAHAYGIKVIADIVINHRRAGDLEWN 145  
122 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 181  
146 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 205  
182 ESYAAYLRSIGDAMRFDYKGYAPWVVDLWLNWGGWAVGEYDNTDNLWVAYSSGA 241  
206 ESYAAYLRSIGDAMRFDYKGYAPWVVDLWLNWGGWAVGEYDNTDNLWVAYSSGA 265  
242 KVFDFALYKMDAEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 301  
266 KVFDFALYKMDAEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 325  
302 AFILTYEGOPTIFRYDYEELNKKLNLWIHNLHAGSGTGGYDPDICHDKSWDQYWLWASQ 361  
326 AFILTYEGOPTIFRYDYEELNKKLNLWIHNLHAGSGTGGYDPDICHDKSWDQYWLWASQ 385  
362 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKXVYSSGWVYLEAPYDPA 421  
386 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKXVYSSGWVYLEAPYDPA 445  
422 NGQYGSVWSYCGVG 436  
446 NGQYGSVWSYCGVG 460

RESULT 8  
Q8U319 PRELIMINARY; PRT; 473 AA.  
AC Q8U319;  
DT 01-JUN-2002 (TremBLrel. 21, Created)  
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)  
DE Alpha-amylase.  
GN PF0477.  
OS Pyrococcus furiosus.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OX Pyrococcus.  
OX NCBI\_TaxID=2261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;  
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
RT "The complete sequence of the Pyrococcus furiosus genome.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE010170; AAL80601.1;  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amy\_cat.  
DR InterPro; IPR006046; Glyco\_hydro\_13.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PRINTS; PR00110; ALPHAAMYLASE.  
KW Complete proteome.  
SQ SEQUENCE 473 AA; 54394 MW; 2E212A2C9A67F3B5 CRC64;

Query Match 87.6%; Score 2153; DB 17; Length 473;  
Best Local Similarity 86.7%; Pred. No. 1.6e-137;  
Matches 377; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

2 AKYLEEGGVIMQAFYWDVPSGGIWDTRIROPKIPWYDAGISAIWIPPSKMGSGGYSM 61  
26 AKYLEEGGVIMQAFYWDVPSGGIWDTRIROPKIPWYDAGISAIWIPPSKMGSGGYSM 85  
62 GYDPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEWN 121  
86 GYDPDYDFDLGEYDQKGTETRFSGKQELVRLIQTAAHAYGIKVIADIVINHRRAGDLEWN 145  
122 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 181  
146 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 205

Db 39 AKYLEEGGVIMQAFYWDVPSGGIWDTRIROPKIPWYDAGISAIWIPPSKMGSGGYSM 98  
QY 62 GYDPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEWN 121  
Db 99 GYDPDYDFDLGEYDQKGTETRFSGKQELVRLIQTAAHAYGIKVIADIVINHRRAGDLEWN 158  
QY 122 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 181  
Db 159 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 218  
QY 182 ESYAAYLRSIGDAMRFDYKGYAPWVVDLWLNWGGWAVGEYDNTDNLWVAYSSGA 241  
Db 219 ESYAAYLRSIGDAMRFDYKGYAPWVVDLWLNWGGWAVGEYDNTDNLWVAYSSGA 278  
QY 242 KVFDFALYKMDAEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 301  
Db 279 KVFDFALYKMDAEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 338  
QY 302 AFILTYEGOPTIFRYDYEELNKKLNLWIHNLHAGSGTGGYDPDICHDKSWDQYWLWASQ 361  
Db 339 AFILTYEGOPTIFRYDYEELNKKLNLWIHNLHAGSGTGGYDPDICHDKSWDQYWLWASQ 398  
QY 362 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKXVYSSGWVYLEAPYDPA 421  
Db 399 KPGLITYINLSSKAGRWVYVPKAGACIHEYTGNLGGWVDKXVYSSGWVYLEAPYDPA 458  
QY 422 NGQYGSVWSYCGVG 436  
Db 459 NGQYGSVWSYCGVG 473

Query Match 87.5%; Score 2152; DB 1; Length 460;  
Best Local Similarity 86.7%; Pred. No. 1.8e-137;  
Matches 377; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

2 AKYLEEGGVIMQAFYWDVPSGGIWDTRIROPKIPWYDAGISAIWIPPSKMGSGGYSM 61  
26 AKYLEEGGVIMQAFYWDVPSGGIWDTRIROPKIPWYDAGISAIWIPPSKMGSGGYSM 85  
62 GYDPDYDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEWN 121  
86 GYDPDYDFDLGEYDQKGTETRFSGKQELVRLIQTAAHAYGIKVIADIVINHRRAGDLEWN 145  
122 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 181  
146 PFVNDYTWTFDFSKVASKGKTYANLDFHFNELHAGSGTGGYDPDICHDKSWDQYWLWASQ 205



```
Db 245 NMTGVDFRMDAVKHLIKYQLQWIDHLRWKTKGELFTVGEYWNVDYNQHLNFIKTS 304
QY 241 AKVDFAL---YKXDEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANEHDTII--- 294
Db 305 MSLFDAPLHNFNFNASKSGNYD---MQIIMNG-TLMKDNFVKAVTLVNEHDTQLOAL 359
QY 295 -----WNKYPAYAFI-LTYEGQPTIFRYDYBEWLNKDK-----LKNLIWIHEN 336
Db 360 ESTVDWFKPLAVAFILLREEGVPSVFYADYGAQYSDKGYNNMAKVPYIELVTLRKE 419
QY 337 LAGGSTDIVYNDDELIFVRNGYDKPGLIYINLGSSKAGRWYVVPKAGACIHEYTG 395
Db 420 YAYGKQN-SYLDHWDVIGWTRGDAAEHFNSMAVIMSDGPGTKMY-----TGKPSTRYVD 474
QY 396 NLG-----GWVDKYVYSSGWVYLEAPAYDPANGQYGVSW 430
Db 475 KLGRTEEVWTD-----ANGW-----AEPVNGG-SVSW 503

RESULT 12
Q8YU21 PRELIMINARY; PRT; 492 AA.
AC Q8YU21,
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE Alpha-amylase.
GN ALR2190.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RL cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003588; BAB73889.1; -.
DR PIR; AH2079; AH2079.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR Pfam; PF00128; alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAMYLASE.
DR Complete Proteome.
KW SEQUENCE 492 AA; 57063 MW; 11B5D6CECF2F18288 CRC64;
SQ SEQUENCE 492 AA; 57063 MW; 11B5D6CECF2F18288 CRC64;

Query Match 20.0%; Score 492; DB 16; Length 492;
Best Local Similarity 27.8%; Pred. No. 3.1e-25;
Matches 147; Conservative 81; Mismatches 153; Indels 148; Gaps 28;

QY 8 EKGVIQAFVWDVPSGIWDTTRQKIPEDYDAGISAIWIPPSKMGAGSYNGYDPYD 67
Db 3 QMGNGMQYFHWYTPNDGNLMSKVEASAPELADAGFTAMWLPYKGFAGSFDVGVGYD 62
QY 68 PFDLGEYDQKGTVTRFGSKQELVNMINTAHYGMKVIADIVNHRAGGDL-----EWNPF 123
Db 63 LFDLGEFDQKSVTKYTRQYLDVAKSLQTHGLQYADAVLNHRKGGDAVETPKATPF 122
QY 124 VND-----YTWDFSEKVASGKYTANY-----LDFHPNELHAGDSGT- 159
Db 123 PQDRLNPKGGLQDIKITYHTNPPG-RQGY-SNFEWHWHFDAVDY-NEYNNGDRSTV 178
QY 160 -----FGGYP-----DICHDKSWDQ-----YWLWASOESYAAVLRISGID 194
Db 179 YLLSGKFNDDYVALEKGNFAYLMGCCDLDFQNEWRGVEVTYWGKWC-----LDTTKVD 230
```

```
QY 195 AWRFDYKGVAPVWVKNLWNGWA-----VEEYDNTVDVNLW-----AYSSGAKVFD 245
Db 231 GFRIDAIKHISTWTFPEWIDALESHACKDLFWGEYWNVDINTLL-WVDVAVRKMVSVD 289
QY 246 FALYKXDEA-----FDKNIPALVSALONGQTVVSRDPFKAVTFVANEHDTII----- 294
Db 290 VFLHYNFHQASKSGNVDMMRI-----LDG-TMMQORPHTAVTFVENHDSQLOALES 341
QY 295 -----WNKYPAYAFI-LTYEGQPTIFRYD-----YEEWLNKDK-----LKNLI 331
Db 342 VVEWFKPLAVAFILLREEGVPCVFHADYGAEDW-GKDGRRYNI FMPSHRWIIDKLL 400
QY 332 WIHENLAGGSTDIVYNDDELIFVRNGYDKPGLIYINLGSSKAGRWYV-----PKPA 386
Db 401 YARKHYAYG-PQVNYLDHWNITGWTRGLGDADHPQGMVIMSDGSEGIKMWVEGKPKTKFI 459
QY 387 GACIHEYTGNLGGVVDKYVYSSGWVYLEAPAYDPANGQY-----GYSVW 430
Db 460 DLTEH-----IKEAVTYNEGW-----GEFCLGGSVSVW 489

RESULT 13
Q8Z5S5 PRELIMINARY; PRT; 494 AA.
AC Q8Z5S5,
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytoplasmic alpha-amylase.
GN SY2171 OR AMYA OR T0914.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Feltwell T., Hamlin N., Haque R.M., Dowd L., White N., Farrar J.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RL enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
Dang W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanlani V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
RT EMBL; AL627272; CAD05711.1; -.
DR EMBL; AS016837; AAO68592.1; -.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp amyl cat_sub.
DR Pfam; PF00128; alpha-amylase; 1.
DR SMART; SM00642; Amy; 1.
DR Complete proteome.
KW SEQUENCE 494 AA; 56451 MW; 5D221A7B24F67D4 CRC64;

Query Match 19.8%; Score 487.5; DB 16; Length 494;
Best Local Similarity 28.7%; Pred. No. 6.2e-25;
Matches 139; Conservative 71; Mismatches 156; Indels 119; Gaps 21;
```

QY 9 KGVIMQAFYMDVPSGGIWDITIRQIPWYDAGISAIWIPSPKMGAGYSGYDYPDF 68  
 DB 2 KNPITLQYFHYFYPDGGKMLSELAEADGLNDGINWVLPACKGASGYSGYDSDL 61  
 QY 69 FDLGYDQKGVETRFSGKQELVNNINTAHAYGMKVIAIDIVINHRAGGDLNMPFV 124  
 DB 62 FDLGFDQKGTIATKYGDKRQLLTAIDALKGNIAVLVDVNNHKGADKEIRIVORVN 121  
 QY 125 -NDYT-----WTFDSKVA-SGKYTANYLDHF-----PHELHAGDSGF- 160  
 DB 122 QDRTQIDNLIIECGWTRTFPARAGQYNSFIWDYHCFSGIDHIENPDE-----DGIFK 176  
 QY 161 -----GGYDICHDK--SWD-----QYWL-WASQESYAAAYLRSIG 192  
 DB 177 IVNDYTGQWNDQVDELGNFDYLMGENIDFRNHAVTEIEIKYARWVMEQTHC----- 229  
 QY 193 IDAWEFDYVKGYPAPWVKDLNMGWGA-----VGEYWDNTVDVNLWAYSSGAK--VP 244  
 DB 230 -DGFRLDAVKHIDPAWFKYKWEHIEHVOAVAPKPLFIVAETWSEHVDKQTYIDQDQKTMDF 288  
 QY 245 DFALYKVMDEA-----FDKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTII----- 294  
 DB 289 DAPLQMKPEASROGAEYDMRHIFT-----GTLVEADPFEAVTLVANHDTQLQALE 340  
 QY 295 -----WNVYPAYAFILTYE-GOPTIEYRD-----EVLNKD-----KLNLIW 332  
 DB 341 APVEWFPPLAYALLILRENGVPSVFPPDLVYGSYEDNGENGTCRVDMPVINQLDRLIL 400  
 QY 333 IHENLAGSTDIVYDNDELI-FVRNGYDQKPLITVINLGSSKAGRWVVPKFAGACIH 391  
 DB 401 ARQRFPAHG-IQTLFDPHPNCIAFSRSGTEENPGCVVVLNMGDDGCKETLLIGDNYANKTW 459  
 QY 392 EYTG 396  
 DB 460 DFLGN 464

RESULT 14  
 Q93148  
 ID Q93148 PRELIMINARY; PRT; 501 AA.  
 AC Q93148;  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Amylase.  
 GN AMYK38.  
 OS Bacillus sp. KSM-K38.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=129736;  
 RN [1]\_TaxID=129736;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KSM-K38;  
 RA Hayashi Y.;  
 RT "Isolation of a new Bacillus alpha-amylase.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB051102; BAB71820.1; -  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha\_amyl\_cat.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 SQ SEQUENCE 501 AA; 57485 MW; 1240F46739A5CC11 CRC64;

Query Match 19.8%; Score 486; DB 2; Length 501;  
 Best Local Similarity 30.28; Pred. No. 8e-25;  
 Matches 149; Conservative 55; Mismatches 163; Indels 126; Gaps 24;

QY 11 GVIMQAFYMDVPSGGIWDITIRQIPWYDAGISAIWIPSPKMGAGYSGYDYPDF 70  
 DB 26 GTMYYEWHLENDQGHNLHDDAAALSAGITAIWIPPAKGNQA-DVGYGAYDLYD 84  
 QY 71 LGEYDQKGVETRFSGKQELVNNINTAHAYGMKVIAIDIVINHRAGGDL----- 118

DB 85 LGFENQKGVTRTKYTKAQLEAIGLSKNSDINVDVNNHKGADFTTEAVQAVQVNP 144  
 QY 119 -EWNPFNDYT---WTFDSKVASGKYTANYLDHPHNLHAGSGTGGYDPICHKSWDQ 174  
 DB 145 NRWODISGAYTIDAWTGFD--FSGRNA--YSDFKRWFEH-----FNGV-----DWDQ 188  
 QY 175 -----YWLWASQESYAAAYLRSIG-----IDA 195  
 DB 189 RYQENHIFRPANTNNWVRDEENGYDYLGNISIDFSHPVEQDELKMGSWFTDELDDG 248  
 QY 196 WRFPYKGYAPWVKDLNMGWGA-----WAVECYWDNTVDV-----LWAYSSGAKV 243  
 DB 249 YRLDAIKHIFWYTSMDWRHNEADODLFVGEYKDKDDVGALEFYLDENMWEMS---L 304  
 QY 244 DFALYKVMDEA-----FDKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTD----- 292  
 DB 305 FDVPLNIFYEASQCGSSYDMRNI-----LRG--SLVEAHPMHAFTFVDNHDTPQGESL 356  
 QY 293 -----IINKYPAFAFILTVE-GOPTIFYDYSEWLN-----KDKLNLWIHENLAGS 341  
 DB 357 ESWVADWFKPLATATILTRGGYPNVFYGYGIPNDNISAKDMIDELLDARQYAYG- 415  
 QY 342 TDIVYDNDELI-FVRNGYDQK--GLITVINLGSSKAGRWVVPK--PAGACIHEVTGNL 397  
 DB 416 TQHDYFDHWDVVGWTRGSSSRFNSGLATIMSGPG--GSKWYVGRQAGQWTDLTGN 474  
 QY 398 GGWDKVVYSSGW 410  
 DB 475 GASVT--INGDGN 485

RESULT 15  
 O82839  
 ID O82839 PRELIMINARY; PRT; 516 AA.  
 AC O82839;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Amylase.  
 OS Bacillus sp.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1409;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KSM-1378;  
 RA MEDLINE=98342096; PubMed=9675143;  
 RA Igashiki K., Hatada Y., Ikawa K., Araki H., Ozawa T., Kobayashi T.,  
 Ozaki K., Ito S.;  
 RT "Improved thermostability of a Bacillus alpha-amylase by deletion of  
 an arginine-glycine residue is caused by enhanced calcium binding.";  
 RL Biochem. Biophys. Res. Commun. 248:372-377(1998).  
 DR EMBL; AB008763; BAA32431.1; -  
 DR HSP; P06278; 1VGS.  
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR006047; Alpha\_amyl\_cat.  
 DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
 DR InterPro; IPR006046; Glyco\_hydro\_13.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PRINTS; PR00110; ALPHAAMYLASE.  
 DR SMART; SM00642; Aamy; 1.  
 SQ SEQUENCE 516 AA; 58841 MW; D90A8C90ECC182F8 CRC64;

Query Match 19.2%; Score 472.5; DB 2; Length 516;  
 Best Local Similarity 29.9%; Pred. No. 6.8e-24;  
 Matches 147; Conservative 66; Mismatches 156; Indels 119; Gaps 25;

QY 11 GVIMQAFYMDVPSGGIWDITIRQIPWYDAGISAIWIPSPKMGAGYSGYDYPDF 70  
 DB 38 GTMYYEWHLENDQGHNLHDDAAALSAGITAIWIPPAKGNQA-DVGYGAYDLYD 96  
 QY 71 LGEYDQKGVETRFSGKQELVNNINTAHAYGMKVIAIDIVINHRAGGDL-----LEWN- 121

Db 97 LGEFNGKGTVRKYGTRSLQCAVTSLKXNGIQVYGVVNMHKGAGDGTENVNAVEVNR 156  
QY 122 ----PFVNDYT---WTFDSKVASGKYTANY-----LDFPNELHAGDSGTFGG--- 162  
Db 157 NRNQBSISGEYTIETAWTKFDFPGRGNTSHSNFKWRWYHFDGTDWDQSRQLQNKIYKFRGTGK 216  
QY 163 -----YFDICH-----KSNQYWLNASQBSYAAYLRSIGIDAWR 197  
Db 217 ANDWEVDIENGNYDILMYADIIMDHPEVINELRNG-----YNTLNLDOGR 265  
QY 198 FDYVKGYAPWVVKDML-----NWGG--WAVEGYWDTNVDVNL-----WAYSSGAKVFD 245  
Db 266 IDAVKHIKYSYTRDMLTHVRNTTGKPMFAVEFWKNDLAAIENYLKNTSWNHS-----VFD 321  
QY 246 FALYYKMDA-----FDKNIPALVSALONGQTVVSRDPPKAVTFVANHTD----- 292  
Db 322 VPLHYNLNNSNGGYFDMRNI-----LNG-SYVQKHPITHAVTFVDNHDSPQGEALES 373  
QY 293 --IIWNKYPAYAFILTYE-GOPTIFVRYDEEWLN-----KDKLKNLIWIHENLAGGSTD 343  
Db 374 FVQSNFKPLAYALILITREGYPSVYGYGCTPHGVPSMKSXIDPILLOARQTYAYG-TQ 432  
QY 344 IVYDNDLI-FVRNGYDKP--GLITYINLGSSKAGRWVYVPKP-AGACIHEVTGNLGG 399  
Db 433 HDYFDHDIIGWTRGDSHPSHNSGLATINSDPG-GNKWMYVGHKAGQVWREDITGNRSG 491  
QY 400 WVDKXVYSSGW 410  
Db 492 TVT--INADGW 500

Search completed: June 29, 2004, 09:35:44  
Job time : 49 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:33:12 ; Search time 22 Seconds  
(without alignments)  
1023.133 Million cell updates/sec

Title: US-10-081-739A-2

Perfect score: 2459

Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGQGYGVNSYCGVG 436

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	488	19.8	485	2	US-08-446-803-2
2	488	19.8	485	2	US-08-861-837-2
3	488	19.8	485	3	US-08-600-656-2
4	488	19.8	485	3	US-09-170-670-2
5	488	19.8	485	3	US-09-170-670-8
6	488	19.8	485	3	US-09-133-068-2
7	488	19.8	485	3	US-09-193-068-8
8	488	19.8	485	3	US-09-183-412-2
9	488	19.8	485	3	US-09-183-412-8
10	488	19.8	485	3	US-09-264-097-5
11	488	19.8	485	3	US-09-354-191A-2
12	488	19.8	485	4	US-09-291-023A-4
13	488	19.8	485	4	US-09-291-023A-14
14	488	19.8	485	4	US-09-290-734-2
15	488	19.8	485	4	US-09-290-734-8
16	488	19.8	485	4	US-09-381-687-3
17	488	19.8	485	4	US-09-545-586-2
18	488	19.8	485	4	US-09-545-586-8
19	488	19.8	485	4	US-09-540-715A-4
20	488	19.8	485	4	US-09-540-715A-14
21	488	19.8	485	4	US-09-769-864-2
22	488	19.8	485	4	US-09-769-864-8
23	486	19.8	501	4	US-09-465-519-4
24	484	19.7	485	4	US-09-291-023A-2
25	484	19.7	485	4	US-09-291-023A-15
26	484	19.7	485	4	US-09-540-715A-2
27	484	19.7	485	4	US-09-540-715A-15

28 478.5 19.5 483 2 US-08-600-908A-13 Sequence 13, Appl  
29 478.5 19.5 483 3 US-08-683-838A-13 Sequence 13, Appl  
30 478.5 19.5 483 4 US-09-636-252A-13 Sequence 13, Appl  
31 475.5 19.3 515 4 US-09-381-687-7 Sequence 7, Appl  
32 474 19.3 501 4 US-09-465-519-2 Sequence 2, Appl  
33 473.5 19.3 560 3 US-08-814-052-6 Sequence 6, Appl  
34 473.5 19.3 560 3 US-08-812-829-6 Sequence 6, Appl  
35 472.5 19.2 480 3 US-09-182-859-4 Sequence 4, Appl  
36 472.5 19.2 480 3 US-09-170-670-5 Sequence 5, Appl  
37 472.5 19.2 480 3 US-09-193-068-5 Sequence 5, Appl  
38 472.5 19.2 480 3 US-09-183-412-5 Sequence 5, Appl  
39 472.5 19.2 480 4 US-09-290-734-5 Sequence 5, Appl  
40 472.5 19.2 480 4 US-09-672-459-4 Sequence 5, Appl  
41 472.5 19.2 480 4 US-09-545-586-5 Sequence 5, Appl  
42 472.5 19.2 480 4 US-10-186-042-4 Sequence 4, Appl  
43 472.5 19.2 480 4 US-09-769-864-5 Sequence 5, Appl  
44 472.5 19.2 483 4 US-09-291-023A-16 Sequence 16, Appl  
45 472.5 19.2 483 4 US-09-537-168-6 Sequence 6, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-446-803-2  
; Sequence 2, Application US/08446803  
; Patent No. 5824531

; GENERAL INFORMATION:

; APPLICANT: Ctrup, Helle

; APPLICANT: Bisgaard-Frantzen, Henrik

; APPLICANT: Ostergaard, Peter Rahbek

; APPLICANT: Rasmussen, Michael Dolberg

; APPLICANT: Van Der Zee, Pia

; TITLE OF INVENTION: Alkaline Bacillus Amylase

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5824531o No. 5824531disk of No. 5824531th America

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,803

; FILING DATE: 01-June-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Harrington, James J.

; REGISTRATION NUMBER: 38,711

; REFERENCE/DOCKET NUMBER: 4157.204-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 867-0123

; TELEFAX: (212) 878-9655

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 485 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-446-803-2

Query Match 19.8%; Score 488; DB 2; Length 485;

Best Local Similarity 31.0%; Fred. No. 1.1e-34;

Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GYIMQAFYWDVPSGGIMWDTIRQKIPYWDAGISAIWIPPSKMGGAYSGYDYPDFD 70



Db 7 GTWQYFBEWHPNDGNHNRRLDDASNLNRGITAIPWPKGT-SQNDVGYGAYDLYD 65  
QY 71 LGEYDQKGTETRGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTETRGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
QY 123 -----FVNDYT---WTFDSKVASKYNTANYLDHFNELHAGDGTGFGYPDICHDKSWDQ 174  
Db 126 NRQEQISGDYITAEATKDFPGRGN---TYSDFKWRWYH-----FDGV-----DWDQ 169  
QY 175 -----YWLWASQES-----YAA-----YLRSIG 192  
Db 170 SRQFQNRKYKFRGDKAWDEVDSENGNYDYLMDYADVMDHPEVNVNLRWGEWYNTLN 229  
QY 193 IDAWRFDYVKGYPWVWVWDL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
Db 230 LDGFRIDAVKHKYSTFRDLWTHVRNATGKEMFAVAEFWKNLGALENYLNKTNWHS-- 287  
QY 241 AKVDFPALLYKDEAFDNKNI PALVSALONGOTVSRDPFKAVTFVANHTD----- 292  
Db 288 --VDFVPLHYNLYNA-SNSGGNYDMAKLLNG-TVQKHPHVAFTFVDNHDSDQGESLESF 343  
QY 293 -IINWKYPAYAFILTYE-GOPTIFRYDY-----EELNKKDLKNIWIHENLAGGSTDI 344  
Db 344 VQEWFKPLAYALLTREQGYESVFGYGIPTHSVPAMKAKIDPILBARQNFAYG-TQH 402  
QY 345 VYDNDLI-FVRNGYDKP--GLITYINLGSKAGRWVYV-PKAGACIHEYTGNLG 400  
Db 403 DYFDHNIIGTREGNTTHNSGLATIMSDGPG-GEKMYVVGONKAGQVWHDITGNKPGT 461  
QY 401 VDKYVSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

## RESULT 2

US-08-861-837-2  
; Sequence 2, Application US/08861837  
; Patent No. 5856164  
; GENERAL INFORMATION:  
; APPLICANT: Ottup, Helle  
; APPLICANT: Bisgard-Frantzen, Henrik  
; APPLICANT: Ostergaard, Peter Rabek  
; APPLICANT: Rasmussen, Michael Dolberg  
; APPLICANT: Van Der Zee, Pia  
; TITLE OF INVENTION: Alkaline Bacillus Amylase  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: No. 58561640 No. 5856164disk of No. 5856164th America  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/08/961.837  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/446.803  
; FILING DATE: 01-June-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harrington, James J.  
; REGISTRATION NUMBER: 38,711  
; REFERENCE/DOCKET NUMBER: 4157.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 867-0123  
; TELEFAX: (212) 878-9655

## ; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 485 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-861-837-2

## Query Match 19.8%; Score 488; DB 2; Length 485;

Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYNDVPSGGIWDITRQKIPEDYDAGISAIWIPPPASKGMGAYSGYDYPDEFD 70  
Db 7 GTWQYFBEWHPNDGNHNRRLDDASNLNRGITAIPWPKGT-SQNDVGYGAYDLYD 65  
QY 71 LGEYDQKGTETRGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTETRGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
QY 123 -----FVNDYT---WTFDSKVASKYNTANYLDHFNELHAGDGTGFGYPDICHDKSWDQ 174  
Db 126 NRQEQISGDYITAEATKDFPGRGN---TYSDFKWRWYH-----FDGV-----DWDQ 169  
QY 175 -----YWLWASQES-----YAA-----YLRSIG 192  
Db 170 SRQFQNRKYKFRGDKAWDEVDSENGNYDYLMDYADVMDHPEVNVNLRWGEWYNTLN 229  
QY 193 IDAWRFDYVKGYPWVWVWDL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
Db 230 LDGFRIDAVKHKYSTFRDLWTHVRNATGKEMFAVAEFWKNLGALENYLNKTNWHS-- 287  
QY 241 AKVDFPALLYKDEAFDNKNI PALVSALONGOTVSRDPFKAVTFVANHTD----- 292  
Db 288 --VDFVPLHYNLYNA-SNSGGNYDMAKLLNG-TVQKHPHVAFTFVDNHDSDQGESLESF 343  
QY 293 -IINWKYPAYAFILTYE-GOPTIFRYDY-----EELNKKDLKNIWIHENLAGGSTDI 344  
Db 344 VQEWFKPLAYALLTREQGYESVFGYGIPTHSVPAMKAKIDPILBARQNFAYG-TQH 402  
QY 345 VYDNDLI-FVRNGYDKP--GLITYINLGSKAGRWVYV-PKAGACIHEYTGNLG 400  
Db 403 DYFDHNIIGTREGNTTHNSGLATIMSDGPG-GEKMYVVGONKAGQVWHDITGNKPGT 461  
QY 401 VDKYVSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

## RESULT 3

US-08-600-656-2  
; Sequence 2, Application US/08600656  
; Patent No. 6093562  
; GENERAL INFORMATION:  
; APPLICANT: Bisgard-Frantzen, Henrik  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; TITLE OF INVENTION: AMYLASE VARIANTS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: No. 60935620 No. 6093562disk of No. 6093562th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/600,656  
 ; FILING DATE: 13-FEB-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lambiris, Elias J.  
 ; REGISTRATION NUMBER: 33,728  
 ; REFERENCE/DOCKET NUMBER: 4318.204-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212 867 0123  
 ; TELEFAX: 212 867 0298  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 485 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-600-656-2

Query Match 19.8%; Score 488; DB 3; Length 485;  
 Best Local Similarity 31.0%; Pred. No. 1.le-34;  
 Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;  
 QY 11 GVIMQAFYVDVPSGGIWMDTIRQKIPWYDAGISAIWIPPSKMGGAISMGYPDPDFD 70  
 Db 7 GTMQYFEWHLNDGHNRLRDDASNLNRGITAIPPAWKGT-SQNDVGYGAYDLID 65  
 QY 71 LGEYDQKGTVETRFSGKQELVNMINTAHAYGMKVIADIVINHRAAGD-----LEWNP- 122  
 Db 66 LGEFNQKGTVRTKYGTRSQLESALHAKKNGVQVYGVVNMHKGADATENVLAVENPN 125  
 QY 123 -----FVNDYT---WTFDFSKVASGKYTANYLDHPHNLHAGSGTFCGYPDICHDKSWDQ 174  
 Db 126 NRNQEISGDYTIKWTKEFDFPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169  
 QY 175 -----YWLWASQES-----YAA-----YLSIG 192  
 Db 170 SRQFQNRIRYKFRGDKAWDEVDSENGNYDLYMADVMDHPEVNVNLRWGEWYNTLN 229  
 QY 193 IDAMRFDYKGYAPVVKDWL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
 Db 230 LDGFRIDAVKHIKYSFTRDWLTHVRNATGEMFAVAFWKNLGDALLENLKNKTWNHS-- 287  
 QY 241 AKVFDFALYKQDEAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHTD----- 292  
 Db 288 --VFDVPLHYNLNA-SNSGGNYDMAKLLNG-TVQKHPHMAVTFVNDHDSQPQGESLEF 343  
 QY 293 -IIWKNYPAYAFILTYE-QOPTIFYRDY-----EWLNKDKLKNLIWIHENLAGGSTDI 344  
 Db 344 VQEWFKPLAYALILTREQQPSVFYGYGIPTHSVPAKAKIDPILARQNFAYG-TQH 402  
 QY 345 VYDNDLEI-FVRNGYDKP--GLITYINLGSSKAGRWVYV-PKFAGACIHEYTGNLGGW 400  
 Db 403 DYFDHNIIGTWREGNTHPNSGLATIMSDGPG-GEKWMYVQGNKAGQVWHDITGNKPGT 461  
 QY 401 VDKYVSSGWYLEAPAYDPANGQYGSVW 430  
 Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 4  
 US-09-170-670-2  
 ; Sequence 2, Application US/09170670  
 ; Patent No. 6187576  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Svendsen, Allan  
 ; APPLICANT: Borchert, Torben  
 ; APPLICANT: Bisgard-Frantzen Henrik  
 ; TITLE OF INVENTION: Alpha-Amylase Mutants  
 ; FILE REFERENCE: 5276.200-US  
 ; CURRENT APPLICATION NUMBER: US/09/170,670  
 ; CURRENT FILING DATE: 1998-10-13  
 ; EARLIER APPLICATION NUMBER: 1172/97

; EARLIER FILING DATE: 1997-10-13  
 ; EARLIER APPLICATION NUMBER: 60/063,306  
 ; EARLIER FILING DATE: 1997-10-28  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 485  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus sp.  
 US-09-170-670-2

Query Match 19.8%; Score 488; DB 3; Length 485;  
 Best Local Similarity 31.0%; Pred. No. 1.le-34;  
 Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;  
 QY 11 GVIMQAFYVDVPSGGIWMDTIRQKIPWYDAGISAIWIPPSKMGGAISMGYPDPDFD 70  
 Db 7 GTMQYFEWHLNDGHNRLRDDASNLNRGITAIPPAWKGT-SQNDVGYGAYDLID 65  
 QY 71 LGEYDQKGTVETRFSGKQELVNMINTAHAYGMKVIADIVINHRAAGD-----LEWNP- 122  
 Db 66 LGEFNQKGTVRTKYGTRSQLESALHAKKNGVQVYGVVNMHKGADATENVLAVENPN 125  
 QY 123 -----FVNDYT---WTFDFSKVASGKYTANYLDHPHNLHAGSGTFCGYPDICHDKSWDQ 174  
 Db 126 NRNQEISGDYTIKWTKEFDFPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169  
 QY 175 -----YWLWASQES-----YAA-----YLSIG 192  
 Db 170 SRQFQNRIRYKFRGDKAWDEVDSENGNYDLYMADVMDHPEVNVNLRWGEWYNTLN 229  
 QY 193 IDAMRFDYKGYAPVVKDWL-----NWGG--WAVEGYWDTNVDV-----LNWYSSG 240  
 Db 230 LDGFRIDAVKHIKYSFTRDWLTHVRNATGEMFAVAFWKNLGDALLENLKNKTWNHS-- 287  
 QY 241 AKVFDFALYKQDEAFDNKNIPALVSALONGQTVVSRDPFKAVTFVANHTD----- 292  
 Db 288 --VFDVPLHYNLNA-SNSGGNYDMAKLLNG-TVQKHPHMAVTFVNDHDSQPQGESLEF 343  
 QY 293 -IIWKNYPAYAFILTYE-QOPTIFYRDY-----EWLNKDKLKNLIWIHENLAGGSTDI 344  
 Db 344 VQEWFKPLAYALILTREQQPSVFYGYGIPTHSVPAKAKIDPILARQNFAYG-TQH 402  
 QY 345 VYDNDLEI-FVRNGYDKP--GLITYINLGSSKAGRWVYV-PKFAGACIHEYTGNLGGW 400  
 Db 403 DYFDHNIIGTWREGNTHPNSGLATIMSDGPG-GEKWMYVQGNKAGQVWHDITGNKPGT 461  
 QY 401 VDKYVSSGWYLEAPAYDPANGQYGSVW 430  
 Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 5  
 US-09-170-670-8  
 ; Sequence 8, Application US/09170670  
 ; Patent No. 6187576  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Svendsen, Allan  
 ; APPLICANT: Borchert, Torben  
 ; APPLICANT: Bisgard-Frantzen Henrik  
 ; TITLE OF INVENTION: Alpha-Amylase Mutants  
 ; FILE REFERENCE: 5276.200-US  
 ; CURRENT APPLICATION NUMBER: US/09/170,670  
 ; CURRENT FILING DATE: 1998-10-13  
 ; EARLIER APPLICATION NUMBER: 1172/97  
 ; EARLIER FILING DATE: 1997-10-13  
 ; EARLIER APPLICATION NUMBER: 60/063,306  
 ; EARLIER FILING DATE: 1997-10-28  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 8  
 ; LENGTH: 485  
 ; TYPE: PRT

```

; ORGANISM: Bacillus sp.
US-09-170-670-8
Query Match      19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGAGYSGMDYDPDFD 70
DB 7 GTMMQYFEWHLFNDGNHNRDDASNLNRGITAIIWIPPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVEFRGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122
DB 66 LGEFNQKGTVRTKYGTRSQLESIAHALKNGQVQVYGVVMMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTDFSKVASGKYTANYLDHFNELHAGDSGTGGYPDICHDKSDQ 174
DB 126 NRNQEISGDYIEAWTKDFPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRSIG 192
DB 170 SRQFQNRYYKFRGDKAWDNEVDSENGYDLYMADVDMHDPVNNELRWGEWYNTLN 229

QY 193 IDAWRFVYKGYAPVWVKWL-----NWMGG--WAVEGYMDTNVAV-----LNWYSSG 240
DB 230 LDGFRIDAVKHKISFTRDWLTHTVRNATGKEMFAVAEFKNDLGALENYLNKTNNHS-- 287

QY 241 AKVDFPALLYKMDFAFNKNI PALVSALONGQTVSRDPFKAVTFVANHTD----- 292
DB 288 --VDFVPLHYNLYNA-SNSGGNYDMAKLLNG-TVQKHPMHAFTFVDNHDSPGESLESF 343

QY 293 -LIWKNYPAYAFILTYE-GOPTIFYRDY-----EWLKDKLKNLIWHENLAGSTDI 344
DB 344 VQEWFKPLAYALILITREQSPSVFYGYIGIPTHSPVPAKAKIDPILBARQNFAYG-TQH 402

QY 345 VYDNDLELI-FVRNGYDKP--GLITYINLGSSKAGRWVYV-PKFAACIHEYTGNLGWS 400
DB 403 DYFDHNIIGWTREGNTHPNSGLATIMSDGPG-GEKMYVGNKAGQVWHDITGNKPGT 461

QY 401 VDKVYSSGWTYLAEPAYDPANGQYGSVW 430
DB 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 6
US-09-193-068-2
; Sequence 2, Application US/09193068
; Patent No. 6197565
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjruliff, S ren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/193,068
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-193-068-2

Query Match      19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGAGYSGMDYDPDFD 70
DB 7 GTMMQYFEWHLFNDGNHNRDDASNLNRGITAIIWIPPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVEFRGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122
DB 66 LGEFNQKGTVRTKYGTRSQLESIAHALKNGQVQVYGVVMMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTDFSKVASGKYTANYLDHFNELHAGDSGTGGYPDICHDKSDQ 174
DB 126 NRNQEISGDYIEAWTKDFPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRSIG 192
DB 170 SRQFQNRYYKFRGDKAWDNEVDSENGYDLYMADVDMHDPVNNELRWGEWYNTLN 229

QY 193 IDAWRFVYKGYAPVWVKWL-----NWMGG--WAVEGYMDTNVAV-----LNWYSSG 240
DB 230 LDGFRIDAVKHKISFTRDWLTHTVRNATGKEMFAVAEFKNDLGALENYLNKTNNHS-- 287

QY 241 AKVDFPALLYKMDFAFNKNI PALVSALONGQTVSRDPFKAVTFVANHTD----- 292
DB 288 --VDFVPLHYNLYNA-SNSGGNYDMAKLLNG-TVQKHPMHAFTFVDNHDSPGESLESF 343

QY 293 -LIWKNYPAYAFILTYE-GOPTIFYRDY-----EWLKDKLKNLIWHENLAGSTDI 344
DB 344 VQEWFKPLAYALILITREQSPSVFYGYIGIPTHSPVPAKAKIDPILBARQNFAYG-TQH 402

QY 345 VYDNDLELI-FVRNGYDKP--GLITYINLGSSKAGRWVYV-PKFAACIHEYTGNLGWS 400
DB 403 DYFDHNIIGWTREGNTHPNSGLATIMSDGPG-GEKMYVGNKAGQVWHDITGNKPGT 461

QY 401 VDKVYSSGWTYLAEPAYDPANGQYGSVW 430
DB 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 7
US-09-193-068-8
; Sequence 8, Application US/09193068
; Patent No. 6197565
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjruliff, S ren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/193,068
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-193-068-8

Query Match      19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGAGYSGMDYDPDFD 70
DB 7 GTMMQYFEWHLFNDGNHNRDDASNLNRGITAIIWIPPAWKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVEFRGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122
DB 66 LGEFNQKGTVRTKYGTRSQLESIAHALKNGQVQVYGVVMMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTDFSKVASGKYTANYLDHFNELHAGDSGTGGYPDICHDKSDQ 174
DB 126 NRNQEISGDYIEAWTKDFPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRSIG 192
DB 170 SRQFQNRYYKFRGDKAWDNEVDSENGYDLYMADVDMHDPVNNELRWGEWYNTLN 229

```

```

QY 193 IDAWRFYVKGYPAPWVKDWL-----NWKGG--WAVEGYWDTNVDV-----LNWAYSSG 240
D 230 LDGFRIDAVKHIKYSFTRDMLTHVRNATGKEMFAVAFWKNDLGALENYLNKTNWHS-- 287
QY 241 AKVDFDALYKMDKDEAFDNKNIPALVSALONGQTVSRDPFKAVTFVANHDTD----- 292
D 288 --VFDVPLHNLNLA-SNSGGNYDMAKLLNG-TVQKHVHVAFTFVDNHDSPGESLESF 343
QY 293 -IWNKYPAYAFILTYE-GOPTIFYRDI-----EWLKDKLKNLWIHENLAGGSTDI 344
D 344 VQEMFPLAVALILITREOGVPSVFYGYGIPHTSVPAKAKIDPILARQNFAYG-TQH 402
QY 345 VYDNDLFI-FVRNGYDGP--GLITYINLGSSKAGRWVYV-PKFAGACIHEVTGNLGGW 400
D 403 DYFDHNIIGWTRGNTTHPNSGLATIMDGP-GEKWMYVGNKAGQVWHDITGNKPGT 461
QY 401 VDKYVSSGVYLEAPAYDPANGQYGSVW 430
D 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 8
US-09-183-412-2
; Sequence 2, Application US/09183412
; Patent No. 6204232
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/183,412
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/064,662
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 60/093,234
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 1240/97
; EARLIER FILING DATE: 1997-10-30
; EARLIER APPLICATION NUMBER: PA 1998 00936
; EARLIER FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-183-412-2

```

```

Query Match 19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIQAFYWDVPSGGIWDITRQKIPENYDAGISAIWIPIPPASKMGAGSYMGYDPYDFD 70
D 7 GTMOMYFEWHPNDGNHNRLEDDASNLNRGITAIPPAWKGT-SQNDVGYGAYDLYD 65
QY 71 LGEYDQKGTVTRFGSKQELVNMINTAHAYGMKVIADIVINHAGD-----LEWNP- 122
D 66 LGEFNQKGTVTRKYGTRSQLESIAIHALKNNGVQVYGDVVMNHKGADATENVLAVEVNP 125
QY 123 -----FVNDYT---WTFDSKVASKYTANYLDHPFNELHAGDSTFGGYDICHDKSWDQ 174
D 126 NRNQEISGDYITIAWTKDFPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169
QY 175 -----YWLNASQES-----YAA-----YLSIG 192
D 170 SRQFNRIYKFRGDKAWDEVDSENGNYDLYMAYDMDHPEVNVNLRWGEWYNTLN 229

```

```

QY 193 IDAWRFYVKGYPAPWVKDWL-----NWKGG--WAVEGYWDTNVDV-----LNWAYSSG 240
D 230 LDGFRIDAVKHIKYSFTRDMLTHVRNATGKEMFAVAFWKNDLGALENYLNKTNWHS-- 287
QY 241 AKVDFDALYKMDKDEAFDNKNIPALVSALONGQTVSRDPFKAVTFVANHDTD----- 292
D 288 --VFDVPLHNLNLA-SNSGGNYDMAKLLNG-TVQKHVHVAFTFVDNHDSPGESLESF 343
QY 293 -IWNKYPAYAFILTYE-GOPTIFYRDI-----EWLKDKLKNLWIHENLAGGSTDI 344
D 344 VQEMFPLAVALILITREOGVPSVFYGYGIPHTSVPAKAKIDPILARQNFAYG-TQH 402
QY 345 VYDNDLFI-FVRNGYDGP--GLITYINLGSSKAGRWVYV-PKFAGACIHEVTGNLGGW 400
D 403 DYFDHNIIGWTRGNTTHPNSGLATIMDGP-GEKWMYVGNKAGQVWHDITGNKPGT 461
QY 401 VDKYVSSGVYLEAPAYDPANGQYGSVW 430
D 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 9
US-09-183-412-8
; Sequence 8, Application US/09183412
; Patent No. 6204232
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/183,412
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/064,662
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 60/093,234
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 1240/97
; EARLIER FILING DATE: 1997-10-30
; EARLIER APPLICATION NUMBER: PA 1998 00936
; EARLIER FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-183-412-8

```

```

Query Match 19.8%; Score 488; DB 3; Length 485;
Best Local Similarity 31.0%; Pred. No. 1.1e-34;
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIQAFYWDVPSGGIWDITRQKIPENYDAGISAIWIPIPPASKMGAGSYMGYDPYDFD 70
D 7 GTMOMYFEWHPNDGNHNRLEDDASNLNRGITAIPPAWKGT-SQNDVGYGAYDLYD 65
QY 71 LGEYDQKGTVTRFGSKQELVNMINTAHAYGMKVIADIVINHAGD-----LEWNP- 122
D 66 LGEFNQKGTVTRKYGTRSQLESIAIHALKNNGVQVYGDVVMNHKGADATENVLAVEVNP 125
QY 123 -----FVNDYT---WTFDSKVASKYTANYLDHPFNELHAGDSTFGGYDICHDKSWDQ 174
D 126 NRNQEISGDYITIAWTKDFPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169
QY 175 -----YWLNASQES-----YAA-----YLSIG 192
D 170 SRQFNRIYKFRGDKAWDEVDSENGNYDLYMAYDMDHPEVNVNLRWGEWYNTLN 229
QY 193 IDAWRFYVKGYPAPWVKDWL-----NWKGG--WAVEGYWDTNVDV-----LNWAYSSG 240

```

Db 230 LDGFRIDAVKHKVSFTFRDMLTHVRNATGKEMFAVAFKNDKALLENYLNKTNWHS-- 287  
Qy 241 AKVDFPALLYKMDKAFDNKNIIPALVSALONGQTVSRDPFKAVTFVANHDTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGGNDYMAKLLNG-TVQKHPHVAFTFVDNHDSPQGESLSF 343  
Qy 293 -IIWKKYPAVAFILTYE-GQTFIYRDY-----EVLNKKDKLNLIWHENLAGSTDI 344  
Db 344 VQEWKPLAYALILTRREGYPSVFGYDYGIPTHSVPMKAKDPILEARQNFAYG-TQH 402  
Qy 345 VYDNDLI-FVRNGYDKP--GLITYINLSSKAGRWVYV-PKFAGACIHEYTGNLGWM 400  
Db 403 DYFDHNIIGWTRGNTHPNSGLATIMSDPG-GEKWMYVGONKAGQVWHDTGNKPGT 461  
Qy 401 VDKYVSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 10  
US-09-264-097-5  
; Sequence 5, Application US/09264097  
; Patent No. 6287826  
; GENERAL INFORMATION:  
; APPLICANT: Hendriksen, Barrie Edmund  
; APPLICANT: Hendriksen, Hanne Vang  
; TITLE OF INVENTION: Enzymatic Preparation of Glucose Syrup  
; TITLE OF INVENTION: From Starch  
; FILE REFERENCE: 5278-200-US  
; CURRENT APPLICATION NUMBER: US/09/264,097  
; EARLIER FILING DATE: 1999-03-08  
; EARLIER FILING DATE: 1998-03-09  
; EARLIER APPLICATION NUMBER: 60/079,209  
; EARLIER FILING DATE: 1998-03-24  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus  
US-09-264-097-5

Query Match 19.8%; Score 488; DB 3; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

Qy 11 GVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGAGYSGYDYPDFD 70  
Db 7 GTMMQYFEWHPNDGNHNRLLDASNLNRGITAIWIPPAWKT-SQNDVGYGAYDLYD 65  
Qy 71 LGEYDQKGTVTRFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVTKGTRSQLESALHAKNGVQVYGVDMVNHKGGADATENVLAVENPN 125  
Qy 123 -----FVNDYT---WTDFSKVASGYTANYLDLPHNELHAGDSGTFGGYDICHDKSMDQ 174  
Db 126 NRNOEISGDYTIETATKDFPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169  
Qy 175 -----YVLWASQES-----YAA-----YLRSTG 192  
Db 170 SQFQNRILYKPRGDKADWDEVSNGNYDLYADVDMHDPVNVNELRWGEWYNTNLN 229  
Qy 193 IDAWRFDYKGYAPWVYKDWL-----NWGG--WAVEGYDNTVNAV-----LNWYSSG 240  
Db 230 LDGFRIDAVKHKVSFTFRDMLTHVRNATGKEMFAVAFKNDKALLENYLNKTNWHS-- 287  
Qy 241 AKVDFPALLYKMDKAFDNKNIIPALVSALONGQTVSRDPFKAVTFVANHDTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGGNDYMAKLLNG-TVQKHPHVAFTFVDNHDSPQGESLSF 343  
Qy 293 -IIWKKYPAVAFILTYE-GQTFIYRDY-----EVLNKKDKLNLIWHENLAGSTDI 344

Db 344 VQEWKPLAYALILTRREGYPSVFGYDYGIPTHSVPMKAKIDPILEARQNFAYG-TQH 402  
Qy 345 VYDNDLI-FVRNGYDKP--GLITYINLSSKAGRWVYV-PKFAGACIHEYTGNLGWM 400  
Db 403 DYFDHNIIGWTRGNTHPNSGLATIMSDPG-GEKWMYVGONKAGQVWHDTGNKPGT 461  
Qy 401 VDKYVSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSIW 482

RESULT 11  
US-09-354-191A-2  
; Sequence 2, Application US/09354191A  
; Patent No. 6297038  
; GENERAL INFORMATION:  
; APPLICANT: Bisgard-Frantzen, Henrik  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; TITLE OF INVENTION: AMYLASE VARIANTS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 62970380 No. 6297038disk of No. 6297038th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/354,191A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/600,656  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4318.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 867 0123  
; TELEFAX: 212 867 0298  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 485 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
US-09-354-191A-2

Query Match 19.8%; Score 488; DB 3; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

Qy 11 GVIMQAFVWDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGAGYSGYDYPDFD 70  
Db 7 GTMMQYFEWHPNDGNHNRLLDASNLNRGITAIWIPPAWKT-SQNDVGYGAYDLYD 65  
Qy 71 LGEYDQKGTVTRFGSKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVTKGTRSQLESALHAKNGVQVYGVDMVNHKGGADATENVLAVENPN 125  
Qy 123 -----FVNDYT---WTDFSKVASGYTANYLDLPHNELHAGDSGTFGGYDICHDKSMDQ 174  
Db 126 NRNOEISGDYTIETATKDFPGRGN---TYSDFKRWYH-----PDGV-----DWDQ 169

QY 175 -----YMLWASQES-----YAA-----YLSIG 192  
DB 170 SRQFNRIYKFRGDKAMDEVDSENGNYDLYMADVDMDHPEVNVNELRRGWETNLN 229  
QY 193 IDAWREFDVVKGAPVWVXDL-----NMGG--WAYGEYDNTNDAV-----LNWAYSSG 240  
DB 230 LGFRIDAVKHIKYFTDNLTHVRNATGKEMFAVAEFWKNDLGALENYLKNKTWNHS-- 287  
QY 241 AKVFDPALYYKMDAEPDNKNIPALVSALQNGOTVVSRRDPFKAVTFVANHDTD----- 292  
DB 288 --VFVPLHYNLYNA--SNSGGNYDMKLLNG--TVQKHPEHVAITFVDNHDQPGESLESF 343  
QY 293 --LIWKNYPAYAFILTYE--GQPTIFYRDI-----EWLNKDKLNIWIHENLAGGSTDI 344  
DB 344 VOEWFKPLAYALILIREQGYPSVFYGYGIPTHSVPAKAKIDPILFARQNFAYG--TQH 402  
QY 345 VYDNDDELII-FVRNGYDKP--GLITYINLGSSKAGRWVYV--PKFAGACIHEYTGNLGGW 400  
DB 403 DYFDHNNIIGWTREGNTTHNSGLATIMSDGPG--GEKMYVQGNKAGQVWHDITGNKPGT 461  
QY 401 VDKYVSSGWVLEAPAYDPDPANGQGYGSVM 430  
DB 462 VT--INADGW-----ANFSVNGG--SVSIV 482

## RESULT 12

US-09-291-023A-4  
; Sequence 4, Application US/09291023A  
; Patent No. 6309871  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Borchert, Torben  
; APPLICANT: Nielsen, Bjarne  
; APPLICANT: Nielsen, Vibeke  
; APPLICANT: Hoeck, Lisbeth  
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nucleic A  
; FILE REFERENCE: 5821.010-US  
; CURRENT APPLICATION NUMBER: US/09/291,023A  
; CURRENT FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: DK 1999 00438  
; PRIOR FILING DATE: 1999-03-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-291-023A-4

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 30.5%; Pred. No. 1.1e-34;  
Matches 154; Conservative 68; Mismatches 169; Indels 114; Gaps 24;

QY 11 GVIMQAFYWDVPSGGIWWDTTIRKIPEDWDAGISAIWIPIPPASKMGAYSGYDPPYDFPD 70  
DB 7 GTIMQYFEWNVDPNGQHWNLHNNNAQNLIKAGITAIWIPIPAWKT--SQNDVGYGAYDLYD 65  
QY 71 LGEYDOKGTVEITRFSGKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP 123  
DB 66 LGEFNQKGTVRTKYGTAKELERAIKSLKANGIQVYGVVMMHKGADFTERVQAVEVNPQ 125  
QY 124 VNDY-----TWTFDSKVASGKYTANY-----LDFHPNELHAGDSGTFGGYPD 165  
DB 126 NRNQSVSGTYQIEAWTGFNPGRGHQHSSFKRWYHFDGTDQDQSRQLANRIYKFRG--- 182  
QY 166 ICHDKSWD-----QYWLWASQES-----YAYLSRIGIDAWRFDYVKG 203  
DB 183 --DGRAMDWEVDTEGNYDLYMADVDMDHPEVINELNRGWVYANTLNLDGFRDLAVKH 240  
QY 204 YAPVWVKDNLNWGG-----WAYGEYDNTNDAV-----LNWAYSSGAKVDFEALYYK 251  
DB 241 IKFSFMRDLGHRVGTGKQLFAVAETWKNDLGALENYLSKTNWTMSA-----FDVPLHYN 296

QY 252 MDEA-----FDKNKIPALVSALQNGOTVVSRRDPFKAVTFVANHDTD-----IWN 296  
DB 297 LYQASNSGNYDMRN-----LLNG--TLVQRHPSHAVTFVDNHDTPQGEALESFVQGW 348  
QY 297 KYPAYAFILTYE--GQPTIFYRDI-----KDKLKNLIWIHENLAGGSTDIYVDN 349  
DB 349 KPLAYATILIREQGYPSVFYGYGIPSDGPSYRQCIDPDLKARQQYAYGRQH--DYFDH 407  
QY 350 DELII-FVRNGYDKP--GLITYINLGSSKAGRWVYV--PKFAGACIHEYTGNLGGWVYV 405  
DB 408 WDWIGTREGNASHPSGLATIMSDGPG--GSKMYVGRKAGEVWHDITGNRSGITV--I 464  
QY 406 YSSGWVLEAPAYDPDPANGQGYGSVM 430  
DB 465 NQDQGWGHEF-----VNGG--SVSIV 482

## RESULT 13

US-09-291-023A-14  
; Sequence 14, Application US/09291023A  
; Patent No. 6309871  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Borchert, Torben  
; APPLICANT: Nielsen, Bjarne  
; APPLICANT: Nielsen, Vibeke  
; APPLICANT: Hoeck, Lisbeth  
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nucleic A  
; FILE REFERENCE: 5821.010-US  
; CURRENT APPLICATION NUMBER: US/09/291,023A  
; CURRENT FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: DK 1999 00438  
; PRIOR FILING DATE: 1999-03-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus  
US-09-291-023A-14

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 30.5%; Pred. No. 1.1e-34;  
Matches 154; Conservative 68; Mismatches 169; Indels 114; Gaps 24;

QY 11 GVIMQAFYWDVPSGGIWWDTTIRKIPEDWDAGISAIWIPIPPASKMGAYSGYDPPYDFPD 70  
DB 7 GTIMQYFEWNVDPNGQHWNLHNNNAQNLIKAGITAIWIPIPAWKT--SQNDVGYGAYDLYD 65  
QY 71 LGEYDOKGTVEITRFSGKQELVNMINTAHAYGMKVIADIVINHRAGD-----LEWNP 123  
DB 66 LGEFNQKGTVRTKYGTAKELERAIKSLKANGIQVYGVVMMHKGADFTERVQAVEVNPQ 125  
QY 124 VNDY-----TWTFDSKVASGKYTANY-----LDFHPNELHAGDSGTFGGYPD 165  
DB 126 NRNQSVSGTYQIEAWTGFNPGRGHQHSSFKRWYHFDGTDQDQSRQLANRIYKFRG--- 182  
QY 166 ICHDKSWD-----QYWLWASQES-----YAYLSRIGIDAWRFDYVKG 203  
DB 183 --DGRAMDWEVDTEGNYDLYMADVDMDHPEVINELNRGWVYANTLNLDGFRDLAVKH 240  
QY 204 YAPVWVKDNLNWGG-----WAYGEYDNTNDAV-----LNWAYSSGAKVDFEALYYK 251  
DB 241 IKFSFMRDLGHRVGTGKQLFAVAETWKNDLGALENYLSKTNWTMSA-----FDVPLHYN 296  
QY 252 MDEA-----FDKNKIPALVSALQNGOTVVSRRDPFKAVTFVANHDTD-----IWN 296  
DB 297 LYQASNSGNYDMRN-----LLNG--TLVQRHPSHAVTFVDNHDTPQGEALESFVQGW 348  
QY 297 KYPAYAFILTYE--GQPTIFYRDI-----KDKLKNLIWIHENLAGGSTDIYVDN 349

Db 349 KPLATATILTRQGYQVYFYGYGIPSDGVPYRQCIDPLLKARQQVAYGRQH-DYFDH 407  
QY 350 DELI-FVRNGYDKP--GLITVINLSSKAGRWVVPK-FAGACIHEYTNLGNWYKYV 405  
Db 408 WDVIWGTREGNASHPNGLATIMSDPG--GSKWYVGRQKAGEVWHDMTGNRSGT--I 464  
QY 406 YSSGWVYLEAPAYDPANGQYGSVW 430  
Db 465 NODGMCHFF-----VNGG-SVSVW 482

RESULT 14  
US-09-290-734-2  
; Sequence 2, Application US/09290734  
; Patent No. 6361989  
; GENERAL INFORMATION:  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; APPLICANT: Bisgard-Frantzen Henrik  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Nielsen, Bjarne Ronfeldt  
; APPLICANT: Nielsen, Vibeke Skovgaard  
; APPLICANT: Hoeck, Lisbeth Hedegaard  
; TITLE OF INVENTION: No. 6361989el -Amylase And -Amylase Mutants  
; FILE REFERENCE: 5276.400-US  
; CURRENT APPLICATION NUMBER: US/09/290,734  
; CURRENT FILING DATE: 1999-04-13  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-290-734-2

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYMDVPSSGIWMDTIROKIPEDWDAGISAIWIPPPASKMGAGYSGMDYDYPDFD 70  
Db 7 GTMWMQFEWHLNDGNHNRRLDDASNLNRGITAIPPAWKGKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVETRFPSKQELVNMINTAHAYGMKVIADIVINHAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVRTKYGTRSQLESIAHALKNGVQVGYGVVNMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTFDSKVASGKYTANYLDHPNELHAGDSGTGGYDPDICHDKSWDQ 174  
Db 126 NRNOEISGDYITIEAWTKFDPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRISG 192  
Db 170 SRQFONRIYKFRGDKANDVEVDSENGNYDLYADVDMDPHPEVNMELRWGEWYNTLN 229

QY 193 IDAWRPDYKGYAPWVVKDWL---NWGG--WAVEGYWDTNVDV-----LNWAYSSG 240  
Db 230 LDGFRIDAVKHIIKYSFTRDWTHTVRNATGKEMFAVAFWKNLGALENLYNKTNWHS-- 287

QY 241 AKVDFDFALYKMDFAFNKNIIPALVSALONGQTVSRDPFKAVTFVANHTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGNYDMAKLNG--TWQKHPMAVTFVDNHDSDPGESLESF 343

QY 293 -IWNKYPAYAFILTYE-GOPTIFYRDY-----EWLKDKLKNLIWIHENLAGSTDI 344  
Db 344 VQEFKFLAYALILTRQGYPSVFGYGYGIPTHSVPMKAKIDPILARQNFAYG-TQH 402

QY 345 VYDNDDEL-I-FVRNGYDKP--GLITVINLSSKAGRWVVPK-FAGACIHEYTNLGNW 400  
Db 403 DYFDHNNIIGTREGNTHPNGLATIMSDPG--GKXWMTVGQNKAGQVWHDITGNKPGT 461

QY 401 VDKVYSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSVW 482

Db 462 VT--INADGW-----ANFSVNGG-SVSVW 482

RESULT 15  
US-09-290-734-8  
; Sequence 8, Application US/09290734  
; Patent No. 6361989  
; GENERAL INFORMATION:  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Borchert, Torben Vedel  
; APPLICANT: Bisgard-Frantzen Henrik  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Nielsen, Bjarne Ronfeldt  
; APPLICANT: Nielsen, Vibeke Skovgaard  
; APPLICANT: Hoeck, Lisbeth Hedegaard  
; TITLE OF INVENTION: No. 6361989el -Amylase And -Amylase Mutants  
; FILE REFERENCE: 5276.400-US  
; CURRENT APPLICATION NUMBER: US/09/290,734  
; CURRENT FILING DATE: 1999-04-13  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-290-734-8

Query Match 19.8%; Score 488; DB 4; Length 485;  
Best Local Similarity 31.0%; Pred. No. 1.1e-34;  
Matches 158; Conservative 59; Mismatches 169; Indels 124; Gaps 27;

QY 11 GVIMQAFYMDVPSSGIWMDTIROKIPEDWDAGISAIWIPPPASKMGAGYSGMDYDYPDFD 70  
Db 7 GTMWMQFEWHLNDGNHNRRLDDASNLNRGITAIPPAWKGKGT-SQNDVGYGAYDLYD 65

QY 71 LGEYDQKGTVETRFPSKQELVNMINTAHAYGMKVIADIVINHAGD-----LEWNP- 122  
Db 66 LGEFNQKGTVRTKYGTRSQLESIAHALKNGVQVGYGVVNMHKGADATENVLAVEVNP 125

QY 123 -----FVNDYT---WTFDSKVASGKYTANYLDHPNELHAGDSGTGGYDPDICHDKSWDQ 174  
Db 126 NRNOEISGDYITIEAWTKFDPGRGN---TYSDFKRWYH-----FDGV-----DWDQ 169

QY 175 -----YWLWASQES-----YAA-----YLRISG 192  
Db 170 SRQFONRIYKFRGDKANDVEVDSENGNYDLYADVDMDPHPEVNMELRWGEWYNTLN 229

QY 193 IDAWRPDYKGYAPWVVKDWL---NWGG--WAVEGYWDTNVDV-----LNWAYSSG 240  
Db 230 LDGFRIDAVKHIIKYSFTRDWTHTVRNATGKEMFAVAFWKNLGALENLYNKTNWHS-- 287

QY 241 AKVDFDFALYKMDFAFNKNIIPALVSALONGQTVSRDPFKAVTFVANHTD----- 292  
Db 288 --VFDVPLHYNLYNA-SNSGNYDMAKLNG--TWQKHPMAVTFVDNHDSDPGESLESF 343

QY 293 -IWNKYPAYAFILTYE-GOPTIFYRDY-----EWLKDKLKNLIWIHENLAGSTDI 344  
Db 344 VQEFKFLAYALILTRQGYPSVFGYGYGIPTHSVPMKAKIDPILARQNFAYG-TQH 402

QY 345 VYDNDDEL-I-FVRNGYDKP--GLITVINLSSKAGRWVVPK-FAGACIHEYTNLGNW 400  
Db 403 DYFDHNNIIGTREGNTHPNGLATIMSDPG--GKXWMTVGQNKAGQVWHDITGNKPGT 461

QY 401 VDKVYSSGWVYLEAPAYDPANGQYGSVW 430  
Db 462 VT--INADGW-----ANFSVNGG-SVSVW 482



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 09:35:47 ; Search time 49 Seconds  
(without alignments)  
2515.500 Million cell updates/sec

Title: US-10-081-739A-2  
Perfect score: 2459  
Sequence: 1 MAKYSELEKGGVIMQAFYWD.....AYDPANGGYGVSWNSYGVG 436

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2459	100.0	436	14	US-10-146-662-2
2	2459	100.0	436	14	US-10-105-733-2
3	2459	100.0	436	14	US-10-081-739A-2
4	2424	98.6	436	14	US-10-081-872-72
5	2424	98.6	436	14	US-10-385-305-72
6	2406	97.8	436	14	US-10-081-872-6
7	2406	97.8	436	14	US-10-385-305-6
8	2378	96.7	436	14	US-10-081-872-60
9	2378	96.7	436	14	US-10-385-305-60
10	2377	96.7	436	14	US-10-081-872-38
11	2377	96.7	436	14	US-10-385-305-38
12	2371	96.4	436	14	US-10-081-872-30
13	2371	96.4	436	14	US-10-385-305-30
14	2368	96.3	436	14	US-10-081-872-4
15	2368	96.3	436	14	US-10-385-305-4

16	2360	96.0	436	14	US-10-081-872-40	Sequence 40, Appl
17	2360	96.0	436	15	US-10-385-305-40	Sequence 40, Appl
18	2358	95.9	436	14	US-10-081-872-16	Sequence 16, Appl
19	2358	95.9	436	14	US-10-081-872-62	Sequence 62, Appl
20	2358	95.9	436	15	US-10-385-305-16	Sequence 16, Appl
21	2358	95.9	436	15	US-10-385-305-62	Sequence 62, Appl
22	2356	95.8	436	14	US-10-081-872-20	Sequence 20, Appl
23	2356	95.8	436	14	US-10-081-872-32	Sequence 32, Appl
24	2356	95.8	436	15	US-10-385-305-20	Sequence 20, Appl
25	2356	95.8	436	15	US-10-385-305-32	Sequence 32, Appl
26	2352	95.6	436	14	US-10-081-872-56	Sequence 56, Appl
27	2352	95.6	436	15	US-10-385-305-56	Sequence 56, Appl
28	2350	95.6	436	14	US-10-081-872-10	Sequence 10, Appl
29	2350	95.6	436	14	US-10-081-872-28	Sequence 28, Appl
30	2350	95.6	436	15	US-10-385-305-10	Sequence 10, Appl
31	2350	95.6	436	15	US-10-385-305-28	Sequence 28, Appl
32	2349	95.5	436	14	US-10-081-872-42	Sequence 42, Appl
33	2349	95.5	436	14	US-10-081-872-46	Sequence 46, Appl
34	2349	95.5	436	14	US-10-228-063-1	Sequence 1, Appl
35	2349	95.5	436	15	US-10-385-305-42	Sequence 42, Appl
36	2349	95.5	436	15	US-10-385-305-46	Sequence 46, Appl
37	2349	95.5	518	14	US-10-228-063-15	Sequence 15, Appl
38	2349	95.5	741	14	US-10-228-063-10	Sequence 10, Appl
39	2349	95.5	820	14	US-10-228-063-16	Sequence 16, Appl
40	2347	95.4	436	14	US-10-081-872-14	Sequence 14, Appl
41	2347	95.4	436	15	US-10-385-305-14	Sequence 14, Appl
42	2344	95.3	454	14	US-10-228-063-13	Sequence 13, Appl
43	2344	95.3	460	14	US-10-228-063-14	Sequence 14, Appl
44	2344	95.3	460	14	US-10-228-063-33	Sequence 33, Appl
45	2344	95.3	460	14	US-10-228-063-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1  
US-10-146-662-2  
; Sequence 2, Application US/10146662  
; Publication No. US20030013172A1  
; GENERAL INFORMATION:  
; APPLICANT: Grendash, Joel  
; TITLE OF INVENTION: NOVEL METHODS OF ENZYME PURIFICATION  
; FILE REFERENCE: 09010-109001  
; CURRENT APPLICATION NUMBER: US/10/146,662  
; CURRENT FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: US 60/291,122  
; PRIOR FILING DATE: 2001-05-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 436  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated polypeptide  
US-10-146-662-2

Query Match	100.0%	Score 2459;	DB 14;	Length 436;
Best Local Similarity	100.0%	Pred. No. 1.8e-219;		
Matches 436;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAKYSELEKGGVIMQAFYWDVPSGGIWWDTIRQKIPENYDAGISAIWIPPSKMGGGAYS	60	
Db	1	MAKYSELEKGGVIMQAFYWDVPSGGIWWDTIRQKIPENYDAGISAIWIPPSKMGGGAYS	60	
QY	61	MGYDYPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVITADIVINHRAGDLEW	120	
Db	61	MGYDYPDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVITADIVINHRAGDLEW	120	
QY	121	NPFFVNDYTWTFPSKVASKYANYLDLPHNELHAGDSGTFGGYDPDICHDKSWDQYWLWAS	180	
Db	121	NPFFVNDYTWTFPSKVASKYANYLDLPHNELHAGDSGTFGGYDPDICHDKSWDQYWLWAS	180	

us-10-081-739a-2.rapb

Wed Jun 30 10:27:02 2004

```

QY 181 QESYAALRSIGIDAWRFDYKGYAPVWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240
DB 181 QESYAALRSIGIDAWRFDYKGYAPVWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240
QY 241 AKVDFPALLYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
DB 241 AKVDFPALLYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
QY 301 YAFILTYEGOPTIFYRDYEELNWKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
DB 301 YAFILTYEGOPTIFYRDYEELNWKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
QY 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420
DB 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420
QY 421 ANGQYGSVWSYCGVG 436
DB 421 ANGQYGSVWSYCGVG 436

```

```

RESULT 2
US-10-105-733-2
; Sequence 2, Application US/10105733
; Publication No. US20030138786A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Miller, Carl
; APPLICANT: Kazaoka, Martin
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; TITLE OF INVENTION: ENZYMS HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-107002
; CURRENT APPLICATION NUMBER: US/10105,733
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/081,739
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically engineered
US-10-105-733-2

```

```

Query Match 100.0%; Score 2459; DB 14; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.8e-219;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKYSELEKGVINQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPASKMGCGAYS 60
DB 1 MAKYSELEKGVINQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPASKMGCGAYS 60
QY 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
DB 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
QY 121 NPFVNDITWTFDSKVASKYKTANYLDHFNELHAGDSGTGGYDICHDKSWDQYWLWAS 180
DB 121 NPFVNDITWTFDSKVASKYKTANYLDHFNELHAGDSGTGGYDICHDKSWDQYWLWAS 180
QY 181 QESYAALRSIGIDAWRFDYKGYAPVWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240

```

```

DB 181 QESYAALRSIGIDAWRFDYKGYAPVWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240
QY 241 AKVDFPALLYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
DB 241 AKVDFPALLYKMDKDEAFDKNIPALYSALQNGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
QY 301 YAFILTYEGOPTIFYRDYEELNWKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
DB 301 YAFILTYEGOPTIFYRDYEELNWKDKLNIWIHENLAGGSTDIIVYNDDELIFVRNGYG 360
QY 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420
DB 361 DKPGLITYNLGSSKAGRWVYVPKFAACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420
QY 421 ANGQYGSVWSYCGVG 436
DB 421 ANGQYGSVWSYCGVG 436

```

```

Query Match 100.0%; Score 2459; DB 14; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.8e-219;
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKYSELEKGVINQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPASKMGCGAYS 60
DB 1 MAKYSELEKGVINQAFYWDVPSGGIWDITROKIPWYDAGISAIWIPASKMGCGAYS 60
QY 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
DB 61 MGYPDFDFDLGEYDQKGTETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGDLEW 120
QY 121 NPFVNDITWTFDSKVASKYKTANYLDHFNELHAGDSGTGGYDICHDKSWDQYWLWAS 180
DB 121 NPFVNDITWTFDSKVASKYKTANYLDHFNELHAGDSGTGGYDICHDKSWDQYWLWAS 180
QY 181 QESYAALRSIGIDAWRFDYKGYAPVWVKDNLNMGWNAVGEYWDNTVDVNLWAYSSG 240

```



Db 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTD TDI IWKYPA 300  
 QY 301 YAFILTYEGOPTIFYRDYEEWLNKDKLKNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 Db 301 YAFILTYEGOPTIFYRDYEEWLNKDKLKNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 QY 361 DKPLITYINLGSKAGRWYVYVFKFAGACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420  
 Db 361 SKPLITYINLASSKAGRWYVYVFKFAGACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420  
 QY 421 ANGQYGSVWSYCGVG 436  
 Db 421 ANGQYGSVWSYCGVG 436

RESULT 6  
 US-10-081-872-6  
 ; Sequence 6, Application US/10081872  
 ; Publication No. US20030125534A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT APPLICATION NUMBER: US/10/081,872  
 ; CURRENT FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 436  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated polypeptide  
 US-10-081-872-6

Query Match 97.8%; Score 2406; DB 14; Length 436;  
 Best Local Similarity 97.5%; Pred. No. 1.5e-214;  
 Matches 425; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGGVIMQAFYNDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGGAYS 60  
 Db 1 MAKYSELEGGVIMQAFYNDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGGAYS 60  
 QY 61 MGVDYDFDLGEYDOKGTVEYTRFGSKQELVNMINTAHAYGMKVIADIVINHAGGDLW 120  
 Db 61 MGVDYDFDLGEYDOKGTVEYTRFGSKQELVNMINTAHAYGMKVIADIVINHAGGDLW 120  
 QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGSGTGGYPDICHDKSDQYWLWAS 180  
 Db 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGSGTGGYPDICHDKSDQYWLWAS 180  
 QY 181 QESYAAVLRISIGIDANFRDYVKGYAPWVYKDWLNWGGHAGVEYWDVDAVLNWAYSSG 240  
 Db 181 QESYAAVLRISIGIDANFRDYVKGYAPWVYKDWLNWGGHAGVEYWDVDAVLNWAYSSG 240  
 QY 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTD TDI IWKYPA 300  
 Db 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTD TDI IWKYLA 300

QY 301 YAFILTYEGOPTIFYRDYEEWLNKDKLKNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 Db 301 YAFILTYEGOPTIFYRDYEEWLNKDKLKNLIWHENLAGGSTDI VYDNDDELIFVRNGYG 360  
 QY 361 DKPLITYINLGSKAGRWYVYVFKFAGACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420  
 Db 361 DKPLITYINLGSKAGRWYVYVFKFAGACIHEYTGNLGWDVKYVYSSGWVYLEAPAYDP 420  
 QY 421 ANGQYGSVWSYCGVG 436  
 Db 421 ANGQYGSVWSYCGVG 436

RESULT 7  
 US-10-385-305-6  
 ; Sequence 6, Application US/10385305  
 ; Publication No. US20040018607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT APPLICATION NUMBER: US/10/385,305  
 ; CURRENT FILING DATE: 2003-03-06  
 ; PRIOR APPLICATION NUMBER: US/10/081,872  
 ; PRIOR FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 436  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated polypeptide  
 US-10-385-305-6

Query Match 97.8%; Score 2406; DB 15; Length 436;  
 Best Local Similarity 97.5%; Pred. No. 1.5e-214;  
 Matches 425; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGGVIMQAFYNDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGGAYS 60  
 Db 1 MAKYSELEGGVIMQAFYNDVPSGGIWDITRQKIPWYDAGISAIWIPPSKMGGAYS 60  
 QY 61 MGVDYDFDLGEYDOKGTVEYTRFGSKQELVNMINTAHAYGMKVIADIVINHAGGDLW 120  
 Db 61 MGVDYDFDLGEYDOKGTVEYTRFGSKQELVNMINTAHAYGMKVIADIVINHAGGDLW 120  
 QY 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGSGTGGYPDICHDKSDQYWLWAS 180  
 Db 121 NPFVNDYTWTFDFSKVASGKYTANYLDFHPNELHAGSGTGGYPDICHDKSDQYWLWAS 180  
 QY 181 QESYAAVLRISIGIDANFRDYVKGYAPWVYKDWLNWGGHAGVEYWDVDAVLNWAYSSG 240  
 Db 181 QESYAAVLRISIGIDANFRDYVKGYAPWVYKDWLNWGGHAGVEYWDVDAVLNWAYSSG 240  
 QY 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTD TDI IWKYPA 300  
 Db 241 AKVDFDALYKMDAEAFDNKNI PALVSALQNGQT VVS RDPFKAVTFVANHTD TDI IWKYLA 300

[illegible]

```

RESULT 8
US-10-081-872-60
; Sequence 60, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kerovuo, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081.872
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated polypeptide
US-10-081-872-60

```

Query Match	96.7%	Score	2378	DB	14	Length	436
Best Local Similarity	96.1%	Pred. No.	6e-212				
Matches	419	Conservative	8	Mismatches	9	Indels	0
QY	1	MAKYSELEKGGVIMQAFYWDVPSGGIWNDDTTRQKIPEWYDAGISATWIPASKMGWGAYS	60				
Db	1	MAKYLEEGGVIMQAFYWDVPSGGIWNDDTTRQKIPEWYDAGISATWIPASKMGWGAYS	60				
QY	61	MGYDVPYDFDLGEYDQKQVETRFSGKQELVNMINTAHAYGMKVIADIVINHRRAGGDLEW	120				
Db	61	MGYDVPYDFDLGEYDQKQVETRFSGKQELVNMINTAHAYGIVKVIADIVINHRRAGGDLEW	120				
QY	121	NGPVNDYTWDFDSKVASGKYTANYLDPHPNELHAGDSGTFGGVPDICHDKSNDQYWLWAS	180				
Db	121	NGPVNDYTWDFDSKVASGKYTANYLDPHFNEHLHAGDSGTFGGVPDICHDKSNDQYWLWAS	180				
QY	181	QBSYAAYLRSIGIDAWRFDYKGYAPVYVKDMLNMGWGVAUGEYWDNVDAVLNWAYSSG	240				
Db	181	QBSYAAYLRSIGIDAWRFDYKGYAPVYVKDMLNMGWGVAUGEYWDNVDAVLNWAYSSG	240				
QY	241	AKVEFQFALYYKNDFAFDNKNIPALVUSALONGCTVVSBDPKATVFVANHTDIIWNKYFA	300				
Db	241	AKVEFQFALYYKNDFAFDNKNIPALVDAURYQTVVSBDPKATVFVANHTDIIWNKYFA	300				
QY	301	YAFILTYEGQPIFYRDYEEWLNDKLNLIWIHENLAGGSTDITVYDNDDELIFVRNGVG	360				

Db	301	YAFILTYSGQPTIFRYDYEHWLNKDRLLKNLIWHDHLAAGSTDI VYVDNDELIFVNGVG	360
Qy	361	DPFGLLITVINLGSSKAGRWYVPKFAGACIETHYTNLGGWWDKYVYSSGHWYLEAPYDP	420
Db	361	SKPGLITVINLASSKAGRWYVPKFAGACIETHYTNLGGWWDKYVYSSGHWYLEAPYDP	420
Qy	421	ANGQYGYGSWSYCYGVG	436
Db	421	ANGQYGYGSWSYCYGVG	436

```

RESULT 9
US-10-385-305-60
; Sequence 60, Application US/10395305
; Publication NO. US20040018607A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kerovuo, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated polypeptide
US-10-385-305-60

```

Query Match	96.7%	Score 2378;	DB 15;	Length 436;
Best Local Similarity	96.1%;	Pred. No. 6e-212;		
Matches 419;	Conservative 8;	Mismatches 9;	Indels 0;	Gaps 0
Qy	1	MAKYSELKGGVIMQAFYWDVPSGGIWNWDIIROKIPEWYDAGISALWIIPASKMGCGAYS	60	
Db	1	MAKYLEEGGVIMQAFYWDVPSGGIWNWDIIROKIPEWYDAGISALWIIPASKMGCGYS	60	
Qy	61	MGYPDYPDFDLGEXDQKGTVETREFGSKOEILNMINTAHAYGMKVIADIIVINHRAGGDLEW	120	
Db	61	MGYPDYPDFDLGEXYQKGTVETREFGSKOEILNMINTAHAYGIVKIADIIVINHRAGGDLEW	120	
Qy	121	NPFVNDYTWTFDFSKVASGKYTANYLDPHFNLHAGSGTEGGYPDICHKSDWDQYWLWAS	180	
Db	121	NPFVNDYTWTFDFSKVASGKYTANYLDPHFNLHAGSGTEGGYPDICHKSDWDQYWLWAS	180	
Qy	181	QBSYAALRSIGIDAWRFDYVKGYPAPVVDKLNWGMGWA VGWYDTNVDVAINWAYSSG	240	
Db	181	QBSYAALRSIGIDAWRFDYVKGYPAPVVDKLNWGMGWA VGWYDTNVDVAINWAYSSG	240	
Qy	241	AKVDFALYKXKDEAFDNKNIIPALVSA LONGQTVSSEDPKATVFVANHDTDIIINWKYPA	300	
Db	241	AKVDFALYKXKDEAFDNNNIIPALVDALRYGQTVVSRDPPKATVFVANHDTDIIINWKYPA	300	
Qy	301	YAFILTYEGQPTIFYRDYEEELNKDKLNIWIHENLAGGSTDIVYVDNDELIFVRNGVG	360	

Db 301 YAFILTYEGQPTIFPRDYEEWLNKDKLNLIWHDHLAGGSTDIVVYDNDELIFVRNGYG 360  
 QY 361 DKPLITYINLGSSKAGRWVYVPKFAAGACIHEVTGNLGGWVDKYVYSSGWWVYLEAPAYDP 420  
 Db 361 SKPLITYINLASSKAGRWVYVPKFAAGACIHEVTGNLGGWVDKYVYSSGWWVYLEAPAYDP 420  
 QY 421 ANGQYGSVMSYCGVG 436  
 Db 421 ANGQYGSVMSYCGVG 436

RESULT 10  
 US-10-081-872-38  
 ; Sequence 38, Application US/10081872  
 ; Publication No. US20030125534A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT FILING DATE: 2002-02-21  
 ; PRIOR FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: PastSeq for Windows Version 4.0  
 ; SEQ ID NO 38  
 ; LENGTH: 436  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated polypeptide  
 US-10-081-872-38

Query Match 96.7%; Score 2377; DB 14; Length 436;  
 Best Local Similarity 96.8%; Pred. No. 7.4e-212;  
 Matches 422; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGGVIMQAFYNDVPSSGIWWDTRIROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
 Db 1 MAKYLEEEGGVIMQAFYNDVPSSGIWWDTRIROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
 QY 61 MGYDPYDFDLGEYDQKGTVEFRFGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 Db 61 MGYDPYDFDLGEYDQKGTVEFRFGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWTFDSKVASGKYTANYLDFHNLHAGDSGTFGGYPDICDKSDQYWLWAS 180  
 Db 121 NPFVNDYTWTFDSKVASGKYTANYLDFHNLHAGDSGTFGGYPDICDKSDQYWLWAS 180  
 QY 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
 Db 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
 QY 241 AKVFDPAIYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
 Db 241 AKVFDPAIYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGQPTIFPRDYEEWLNKDKLNLIWHDHLAGGSTDIVVYDNDELIFVRNGYG 360  
 Db 301 YAFILTYEGQPTIFPRDYEEWLNKDKLNLIWHDHLAGGSTDIVVYDNDELIFVRNGYG 360

QY 361 DKPLITYINLGSSKAGRWVYVPKFAAGACIHEVTGNLGGWVDKYVYSSGWWVYLEAPAYDP 420  
 Db 361 SKPLITYINLASSKAGRWVYVPKFAAGACIHEVTGNLGGWVDKYVYSSGWWVYLEAPAYDP 420  
 QY 421 ANGQYGSVMSYCGVG 436  
 Db 421 ANGQYGSVMSYCGVG 436

RESULT 11  
 US-10-385-305-38  
 ; Sequence 38, Application US/10385305  
 ; Publication No. US20040018607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Callen, Walter  
 ; APPLICANT: Richardson, Toby  
 ; APPLICANT: Frey, Gerhard  
 ; APPLICANT: Short, Jay M.  
 ; APPLICANT: Mathur, Eric J.  
 ; APPLICANT: Gray, Kevin A.  
 ; APPLICANT: Kerovuo, Janne S.  
 ; APPLICANT: Slupska, Malgorzata  
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY  
 ; FILE REFERENCE: 09010-108001  
 ; CURRENT FILING DATE: 2003-03-06  
 ; PRIOR FILING DATE: 2003-03-06  
 ; PRIOR APPLICATION NUMBER: US/10/081,872  
 ; PRIOR FILING DATE: 2002-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,495  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/270,496  
 ; PRIOR FILING DATE: 2001-02-21  
 ; PRIOR APPLICATION NUMBER: US 60/291,122  
 ; PRIOR FILING DATE: 2001-05-14  
 ; NUMBER OF SEQ ID NOS: 321  
 ; SOFTWARE: PastSeq for Windows Version 4.0  
 ; SEQ ID NO 38  
 ; LENGTH: 436  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetically generated polypeptide  
 US-10-385-305-38

Query Match 96.7%; Score 2377; DB 15; Length 436;  
 Best Local Similarity 96.8%; Pred. No. 7.4e-212;  
 Matches 422; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MAKYSELEKGGVIMQAFYNDVPSSGIWWDTRIROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
 Db 1 MAKYLEEEGGVIMQAFYNDVPSSGIWWDTRIROKIPEDWDAGISAIWIPPSKGMGGAYS 60  
 QY 61 MGYDPYDFDLGEYDQKGTVEFRFGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 Db 61 MGYDPYDFDLGEYDQKGTVEFRFGSKQELVNMINTAHAYMKVIADIVINHRAGDLEW 120  
 QY 121 NPFVNDYTWTFDSKVASGKYTANYLDFHNLHAGDSGTFGGYPDICDKSDQYWLWAS 180  
 Db 121 NPFVNDYTWTFDSKVASGKYTANYLDFHNLHAGDSGTFGGYPDICDKSDQYWLWAS 180  
 QY 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
 Db 181 QESYAAVLRISIGIDAWRFDYVKGYAPWVVKDWLNWGGWAVGEYWDNTVDALNWAYSSG 240  
 QY 241 AKVFDPAIYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
 Db 241 AKVFDPAIYKMDAEAFDNKNI PALVSALQNGQTVWSRDPFKAVTFVANHTDIIWNKYPA 300  
 QY 301 YAFILTYEGQPTIFPRDYEEWLNKDKLNLIWHDHLAGGSTDIVVYDNDELIFVRNGYG 360  
 Db 301 YAFILTYEGQPTIFPRDYEEWLNKDKLNLIWHDHLAGGSTDIVVYDNDELIFVRNGYG 360





```
Db 361 DKPLITYINLGSSKAGRWVYVPKFAACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
QY 421 ANGOYGYSVMSYCGVG 436
Db 421 ANGOYGYSVMSYCGVG 436

RESULT 14
US-10-081-872-4
; Sequence 4, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kervuo, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated polypeptide
US-10-081-872-4

Query Match 96.3%; Score 2368; DB 14; Length 436;
Best Local Similarity 96.1%; Pred. No. 5e-211;
Matches 419; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAKYSLEKGGVIMQAFYVDVPSGG:WMDTIROKIPENYDAGISAIWIPPASKMGMGAYS 60
Db 1 MAKYLEEGGLIMQAFYVDVPMGGI:WMDTIQKIPDASAGISAIWIPPASKMGMGAYS 60

QY 61 MGYDPDFDLGSDYDOKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120
Db 61 MGYDPDFDLGSDYDOKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120

QY 121 NPFVNDYTWDFSKVASGKYTANYLDPHNPHLAGSGTFFGGYDICHDKSDQYWLWAS 180
Db 121 NPFVNDYTWDFSKVASGKYTANYLDPHNPHLAGSGTFFGGYDICHDKSDQYWLWAS 180

QY 181 QESYAAVLRISIGIDAWRFYKGYAPWVKDMLNMGWNAVGEYWDNTVDVNLWAYSSG 240
Db 181 QESYAAVLRISIGIDAWRFYKGYAPWVKDMLNMGWNAVGEYWDNTVDVNLWAYSSG 240

QY 241 AKVDFPALYYKMDAEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
Db 241 AKVDFPALYYKMDAEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300

QY 301 YAFILTVEGQPTIFRYDYEWLNKDKLNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360
Db 301 YAFILTVEGQPTIFRYDYEWLNKDKLNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360

QY 361 DKPLITYINLGSSKAGRWVYVPKFAACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
Db 361 SKPLITYINLASSEAGRWVYVPKFAACIHEVTGNLGGWVDKYVYSSGWVYLEAPAHDP 420
```

```
QY 421 ANGOYGYSVMSYCGVG 436
Db 421 ANGOYGYSVMSYCGVG 436

RESULT 15
US-10-385-305-4
; Sequence 4, Application US/10385305
; Publication No. US20040018607A1
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kervuo, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/385,305
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/081,872
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated polypeptide
US-10-385-305-4

Query Match 96.3%; Score 2368; DB 15; Length 436;
Best Local Similarity 96.1%; Pred. No. 5e-211;
Matches 419; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MAKYSLEKGGVIMQAFYVDVPSGGI:WMDTIROKIPENYDAGISAIWIPPASKMGMGAYS 60
Db 1 MAKYLEEGGLIMQAFYVDVPMGGI:WMDTIQKIPDASAGISAIWIPPASKMGMGAYS 60

QY 61 MGYDPDFDLGSDYDOKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120
Db 61 MGYDPDFDLGSDYDOKGTVEFRFGSKQELVNMINTAHAYGMKVIADIVINHRAGDLEW 120

QY 121 NPFVNDYTWDFSKVASGKYTANYLDPHNPHLAGSGTFFGGYDICHDKSDQYWLWAS 180
Db 121 NPFVNDYTWDFSKVASGKYTANYLDPHNPHLAGSGTFFGGYDICHDKSDQYWLWAS 180

QY 181 QESYAAVLRISIGIDAWRFYKGYAPWVKDMLNMGWNAVGEYWDNTVDVNLWAYSSG 240
Db 181 QESYAAVLRISIGIDAWRFYKGYAPWVKDMLNMGWNAVGEYWDNTVDVNLWAYSSG 240

QY 241 AKVDFPALYYKMDAEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300
Db 241 AKVDFPALYYKMDAEAFDNKNI PALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPA 300

QY 301 YAFILTVEGQPTIFRYDYEWLNKDKLNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360
Db 301 YAFILTVEGQPTIFRYDYEWLNKDKLNLIWHENLAGGSTDIVYDNDDELIFVRNGYG 360

QY 361 DKPLITYINLGSSKAGRWVYVPKFAACIHEVTGNLGGWVDKYVYSSGWVYLEAPAYDP 420
Db 361 SKPLITYINLASSEAGRWVYVPKFAACIHEVTGNLGGWVDKYVYSSGWVYLEAPAHDP 420
```

Qy 421 ANGOYGYSVWSYCGVG 436  
Db 421 ANGYTYGYSVWSYCGVG 436

Search completed: June 29, 2004, 09:41:42  
Job time : 50 secs

